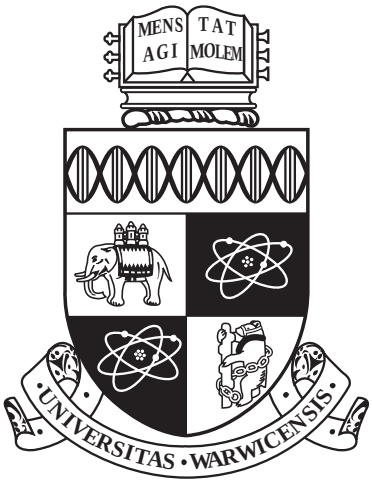



# A novel linkage between maintenance of protein translational fidelity and cell wall biosynthesis in *Streptococcus pneumoniae*

by

**Anna York**

**Thesis**

Submitted to The University of Warwick

for the degree of

**Doctor of Philosophy in Biological Sciences**

**School of Life Sciences**

April 2019

THE UNIVERSITY OF WARWICK

# Contents

# List of Tables

# List of Figures

# Acknowledgments

I would like to express my sincere gratitude to my supervisors Prof David Roper and Dr Adrian Lloyd for their continued support and guidance, and for providing me with the opportunity to work on such an exciting project. I would like to thank Anita Catherwood, Julie Tod and John Moat for their expertise and assistance in the lab, in addition I would like to thank all members of C10 (past and present) for making this such a wonderful experience. I would like to thank my collaborators, Dr. Charo I. del Genio (School of Life Sciences, University of Warwick), Prof Rudo Roemer (Department of Physics, University of Warwick), Prof Syma Khalid and Jonathan Shearer (School of Chemistry, University of Southampton) for their contributions, patience and support. I would also like to thank my advisory panel for their invaluable guidance and support throughout my PhD.

I would also like thank the BBSRC MIBTP for funding and to all the people in my cohort for sharing this journey - it has been a pleasure!

In addition to the above there are a huge number of people whom I would like to acknowledge for their support, guidance, encouragement, friendship, love and kindness who have stood by me and supported me in every way possible throughout this long journey.

I would like to thank all the inspirational teachers and tutors that I was lucky enough to be taught by throughout my time at school and university. Thank you to those who encouraged me to chase my dreams, who helped me up when I fell and believed in me when I didn't believe in myself. I want to thank all the friends I have made over the years, who have continued to stay in touch despite me rarely finding

the time to catch-up or visit. Your friendship means the world to me; Charlotte, Sarah, Emily, Shreya, Adam, Tanweer and Hannah.

I would also like to thank everyone at the University of Warwick who has shared part of this journey with me, made memories that will last a lifetime, provided support and assistance and helped make this experience the wonderful one it has been. Thank you to womens' rugby team, the rowers, the climbers, the Residential Life Team, Disability Support Services, the lab technicians (media prep, wash up and stores), the teaching staff, the admin staff (particularly in the Student Office) and all the various services and support staff at the School of Life Sciences. You have all played an invaluable part in my success thus far, and contribute to making Warwick such a wonderful place to have spent the last 9 years.

The people who have made this PhD something really special, of course, are my supervisors and the people in C10. Thank you, Dave and Adrian, for being the best supervisors I could have hoped for, for being so understanding and for putting up with all the tears. I am incredibly grateful for all of the opportunities you have given me. Thank you for believing in me and supporting me through good times and bad, for giving me so much of your time and sharing your expertise over the years. Thank you for teaching me how to be the best scientist I can be, and for letting me make mistakes and helping me learn from them. Adrian, you have been an inspiration to me ever since I walked into my first undergraduate tutorial with you, it has been a pleasure to work with you. To all the friends I have made in C10, thank you for being like a second family, and making it such a lovely place to work - you are all amazing.

I would also like to say a huge thank you to my special friend Cathy, with whom I have shared emergency tea, recovery hot chocolates, tears of sadness and joy, moments of deepest regret and moments of elation. Having you around has made my time in the lab so much more magical, you have taught me so much and make me want to be better each and every day. I will always treasure the memories we have made. I would also like to thank my wonderful friend Maz who has always been there for me, made me laugh hysterically and encouraged me to not work continuously. Cathy and Maz, you are my closest friends and your mean the world

to me, you inspire me everyday and I simply cannot imagine what this would have been like without you. We have shared this journey from beginning to end and for that I am forever grateful. Thank you to all the other wonderful friends I have made over the last few years, including my totally mad gin buddies Mark, Helen and Eri.

Finally, I would like to thank my wonderful family. My mum (and best friend) who ALWAYS answers the phone (no matter what time of day or night). Thank you for 'walking' me home safely after late nights at the lab and pretending to be interested in science. Mainly though, thank you for being YOU - you have always been an inspiration. Dad, you're the best! Thank you for always believing in me, for being there in the middle of the night when I didn't think I could succeed, and for continually pushing me to be better. Also thanks for all the proofreading and exam preparation you have helped with.

I would also like to apologise to my family; my parents, brothers, sisters-in-law, aunts, uncles and grandparents who have put up with hearing about science way more than you wanted or needed. I love you all, thank you for your support, advice and love, I am truly lucky to have you all. To my beautiful nieces and god daughter, whom I love so very much. The the pictures, videos and updates on your progress have kept me motivated through some of the toughest times - I hope you grow up to love science as much as I do.

Last but not least, thank you to the love of my life, Dom, you are my rock. I can't imagine having done this without you, thank you for everything you have done for me and for the joy and happiness you bring into my life. You have been there through all the ups and downs, we make the best team and I have never had to face anything I didn't want to alone. I dedicate this thesis to you - my love, my life!

# Declarations

The experimental work in thesis is original work conducted by the author, unless stated otherwise. Work was conducted in the School of Life Sciences, University of Warwick between September 2015 and April 2019. All contributions from collaborators are referenced throughout the text.

None of the material presented here has been submitted for a degree previously.

Signed: 

Anna York

Date: 24/04/2019

# Abstract

*Streptococcus pneumoniae* is responsible for a large amount of morbidity and mortality worldwide, and in addition high-level penicillin resistant strains are emerging rapidly. However, MurM is a protein involved in peptidoglycan branching, which may provide hope for the reversal of $\beta$-lactam resistance in pneumococcal infection because it is essential for high-level penicillin resistance. *Streptococcus pneumoniae* produces large amounts of $H_2O_2$ during aerobic metabolism, but lacks the major oxidative stress response mechanisms. Therefore it is proposed that MurM plays an additional role in maintaining translational fidelity during oxidative stress. This work firstly demonstrates that $H_2O_2$ causes a reduction in post-transfer editing of two aminoacyl-tRNA synthetases (AlaRS and ThrRS) *in vitro*, which may result in a decrease in protein synthesis fidelity *in vivo*. Preliminary experiments indicate for the first time, that iGln Lipid II(Lys) can serve as a substrate for MurM in addition to Lipid II(Lys). A range of tRNA substrates have been prepared for investigating the substrate preference of MurM towards mis-charged tRNA. *In silico* modelling of MurM has provided an improved structure, allowing identification of the lipid substrate binding site. Additionally, interactions of MurM with the membrane, Lipid II(Lys) substrate, and membrane phsopholipids have similarly been simulated *in silico*. It is hoped that a better understanding of MurM's structure, mechanism of action, and interplay with other areas of *Streptococcus pneumoniae* metabolism will be important advances, which may assist development of new clinical approaches towards treatment of pneumococcal infection.

# Abbreviations

All abbreviations are used in accordance with the Journal of Biological Chemistry list of general biological and chemical abbreviations and abbreviations of units of measurement and of physical and chemical quantities list. Any additional abbreviations have been detailed below.

- 5P - pentapeptide side chain (commonly L-Ala-$\gamma$-D-Glu-X-D-Ala-D-Ala, whereby X is L-Lys in Gram-positive bacteria or *meso*-DAP in Gram-negative bacteria)
- aa-AMP - aminoacyl-adenylate
- aaRS - aminoacyl-tRNA synthetase
- aa-tRNA - aminoacyl-tRNA (aa can be any amino acid which has been acylated to the tRNA)
- ADPCP - adenosine 5′-(beta-gamma-methylene) triphosphate
- ADPNP - 5′-adenylyl-$\beta$,$\gamma$-imidodiphosphate
- $AgNO_3$ - silver nitrate
- AIM - autoinduction media
- Amp - ampicillin
- ANOVA - analysis of variance
- APS - ammonium persulphate
- AU - absorbance units
- AUC - analytical ultracentrifugation

- BHI - brain heart infusion
- $CaCl_2$ - calcium chloride
- Cam - chloramphenicol
- CAP - Community acquired pneumonise
- CDC- Centers for Disease Control and Prevention
- $CH_3COOH$ - acetic acid
- CID - collision induced dissociation
- CL - cardiolipin
- CLSI - Clinical  Laboratory Standards Institute
- $CO_2$ - carbon dioxide
- CPS - capsular polysaccharide
- CV - column volume
- DAP - diaminopimelic acid
- DEAE - diethylaminoethyl
- DOPE - discrete optimised protein energy
- DTNB - 5,5-dithio-bis-(2-nitrobenzoic acid)
- $e^-$ - electron
- E. coli - *Escherichia coli*
- EF-Tu - elongation factor thermo unstable
- EUCAST - European Committee on Antimicrobial Susceptibility Testing
- eV - electronvolt
- GlcNAc - *N*-acetyl glucosamine
- $^3H$ - tritium (used to indicate amino acids labelled with tritium)
- $H_2O$ - water

- $H_2O_2$ - hydrogen peroxide
- HCl - hydrochloric acid
- HIC - hydrophobic interaction chromatography
- HMM - high molecular mass
- HO• - hydroxyl radical
- HPLC - high performance liquid chromatography
- IDH - isocitrate dehydrogenase
- IDT - Integrated DNA Technologies (Biotech company)
- IMAC - immobilised metal affinity chromatography
- IPD - invasive pneumococcal diseases
- IPP - inorganic pyrophophatase
- IPTG - Isopropyl $\beta$-D-1-thiogalactopyranoside
- IS - insoluble
- ITC - isothermal calorimetry
- iGln - D-iso Glutamine (when used as a prefix to UDP-5P or Lipid II, it indicates the presence of iGln at position two of the pentapeptide)
- Kan - kanamycin
- $k_{cat}$ - turnover number
- $k_{cat}/K_m$ - catalytic efficiency
- KCl - potassium chloride
- $K_m$ - Michaelis-Menten constant (substrate concentration at which the reaction rate is half of $V_{max}$)
- LB - Luria-Bertani
- LBA - Luria-Bertani agar
- LDH - lactate dehydrogenase

- Lipid II(aa) - aa indicates the amino acid present at the third position of the pentapeptide*
- Lipid II(aa)-AA - aa indicated the amino acid which has been appended to the $\epsilon$-amino group of the third position lysine
- LMM - low molecular mass
- LysRS - lysyl-tRNA synthetase
- MBP - maltose binding protein
- MD - molecular dynamics
- MESG - 7-methyl-6-thioguanosine
- $MgCl_2$ - magnesium chloride
- $MgSO_4$ - magnesium sulphate
- MIC - minimum inhibitory concentration
- MDR - multidrug resistance
- $MnCl_2$ - manganese chloride
- MS/MS - tandem mass spectrometry
- MST - microscale thermophoresis
- MurNAc - *N*-acetylmuramic acid
- MW - molecular weight
- MWCO - molecular weight cut-off
- NaCl - sodium chloride
- $Na_2CO_3$ - sodium carbonate
- NaI - sodium iodide
- NaOH - sodium hydroxide
- $Na_2S_2O_3$ - sodium thiosulphate
- NEB - New England Biolabs

- NGS - next generation sequencing
- NMR - nucelar magnetic resonance
- NMT - N-myristoyl transferase
- $O_2$ - oxygen
- OD - optical density
- $OD_{600nm}$ - optical density at 600 nm
- OH - hydroxyl
- PBP - penicillin binding protein
- PDB - protein data bank
- PEP - phophoenolpyruvate
- PG - peptidoglycan
- PhG  phosphatidylglycerol
- PK - pyruvate kinase
- PMSF - phenylmethane sulfonyl fluoride
- PNP - purine nucleoside phosphorylase
- ProRS - prolyl-tRNA synthetase
- psi - pound force per square inch
- $RbCl_2$ - rubidium chloride
- RMSD - root mean square deviation
- ROS - Reactive oxygen species
- SEC - size exclusion chromatography
- SerRS - seryl-tRNA synthetase
- SOAP - statistically optimized atomic potentials
- SOC - super optimal broth with catabolite repression

- SOD - superoxide dismutases
- SolA - solvent A
- SPR - surface plasmon resonance
- T - time
- TC - total count
- TEMED - N,N,N′,N′-Tetramethylethylenediamine
- TEV - Tobacco Etch Virus protease
- TG - transglycosylation
- ThrRS - threonyl-tRNA synthetase
- $T_m$ - melting temperature
- TP - transpeptidation
- $tRNA^{aa}$ - tRNA whereby aa represents the cognate amino acid
- TSB - tryptic soy broth
- *T. thermophilus* - Thermus thermophilus
- U - units (as defined by the manufacturer)
- $V_0$ - initial reaction rate
- Vis visible
- W. viridescens - *Weissella viridescens*

Example structures of starred (*) items can be found in Appendix 7.5.

# Chapter 1

# Introduction

## 1.1 Pathogenesis of *Streptococcus pneumoniae*

*Streptococcus pneumoniae* or pneumococcus is a Gram-positive, non-motile, non spore forming, coccoid bacterium. It most commonly appears as a diplococcus but can also be found individually and in short chains (Toit *et al.*, 2014). The pneumococcus is encapsulated by complex capsular polysaccharides (CPS), which are critical for evasion of the immune system of the host and is an important virulence factor (Geno *et al.*, 2015). The composition of the capsule gives rise to 91 antigenically different serotypes (which belong to 46 serogroups) (Henrichsen, 1995; Park *et al.*, 2007). Streptococci, like other Lactobacillales, are catalase-negative organisms which produce L-(+)-lactic acid as the main product of carbohydrate metabolism (Toit *et al.*, 2014). *S. pneumoniae* is a facultative anaerobe which produces hydrogen peroxide ($H_2O_2$) as a metabolic bi-product during aerobic growth. *Haemophilus influenzae* resides in the same environmental niche, and as such it is possible that the inhibitory and bactericidal effects of $H_2O_2$ may provide a competitive advantage for *S. pneumoniae* (Pericone *et al.*, 2000). The genome of *S. pneumoniae* is relatively small $\sim$2.2 Mbp; whilst a genome of this size would typically reduce the cost of maintaining and replicating a larger genome, its metabolic capacity and ability

to adapt to environmental changes would be limited. However, streptococci are naturally transformable, and so can actively take-up extracellular deoxyribonucleic acid (DNA) and integrate it into their genomes. This provides a mechanism for adjusting to environmental pressures through the acquisition of genes or alleles from a larger shared pool known as the supragenome (Ola *et al.*, 2007). This type of horizontal gene transfer results in huge genetic variation and *S. pneumoniae* is often referred to as having a mosaic genome. Between serotypes, $\sim$10 % of genes are divergent with 160 kb accounting for half of the variable genes, many of which are associated with transposases or are involved in antibiotic resistance (Brückner *et al.*, 2004; Hakenbeck *et al.*, 2001).

*S. pneumoniae* resides asymptomatically in the nasopharynx of many healthy children and adults worldwide. However in susceptible individuals, such as the very young, elderly or immunocomprimised, *S. pneumoniae* is the causative agent of otitis media, community-acquired pneumonia (CAP) and invasive pneumococcal diseases (IPD) such as bacterial meningitis and bacteraemia (Musher *et al.*, 2002). According to the Centers for Disease Control and Prevention (CDC), whilst the incidence of invasive pneumococcal disease (meningitis and bacteraemia) in the United States of America is decreasing, there has been a marginal increase in mortality between 1997 and 2016 (Figure 1.1). Most recent global estimates indicate that 8-12 % of all deaths in children $<$5 years (excluding Human Immunodeficiency Virus (HIV)-positive individuals) are caused by *S. pneumoniae* (O'Brien *et al.*, 2009).

Serotype distribution in the population is constantly fluctuating and depends on many factors including (but not exclusively); age, gender, geographic location, vaccine coverage, antibiotic usage, pre-existing illness and HIV status (Imöhl *et al.*, 2010; Klugman *et al.*, 2004; Cui *et al.*, 2017; Mendez-Lage *et al.*, 2015; Feikin and Klugman, 2002; Geno *et al.*, 2015). However, the severity and outcome of invasive pneumococcal disease is not serotype specific and is instead dependent upon host factors such as age, immunosuppression and underlying disease (Group *et al.*, 2007). In 2000, a 7-valent pneumococcal vaccine was licensed for young children and

studies showed that between 1998 and 2003 the number of hospitalisations due to invasive pneumococcal disease decreased significantly (Shah and Ratner, 2006). The introduction of the vaccine directly reduced the rate of pneumococcal disease in children, with the greatest reduction in children <2 years of age, and conferred an additional indirect effect on the older population due to herd immunity (Whitney *et al.*, 2003; Lexau *et al.*, 2005).

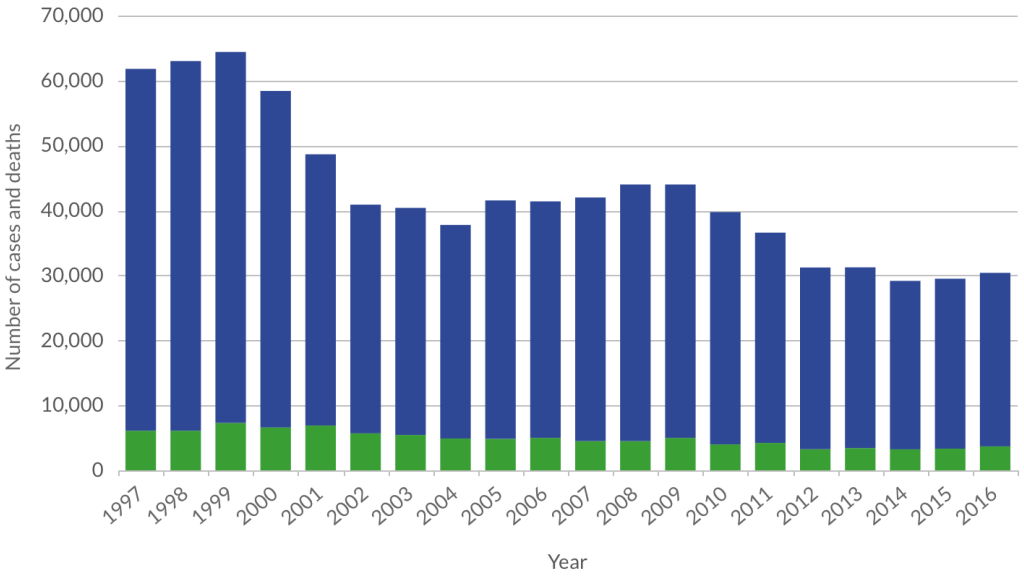



**Figure 1.1: Estimated number of deaths from IPD in the United States of America between 1997 and 2016.** Number of survivals (blue) and number of deaths (green). Image obtained from Centers for Disease Control and Prevention (Centers for Disease Control and Prevention Contributors, 2019).

## 1.2 Antibiotic discovery and the development of resistance

Antibiotics are medicines which inhibit or prevent bacterial growth at low concentrations. A number of different bacteria have been shown to be natural producers of antibiotics and for many years this was believed to provide a competitive advantage against other microorganisms within the same niche (Waksman and Woodruff, 1941; Schatz *et al.*, 1944). However recent studies indicate that antibiotics may act as signalling molecules, regulating homeostasis of

microbial communities by affecting bacterial motility, biofilm formation and protein expression (Linares *et al.*, 2006). There are several historical cases which indicate that humans have exploited and benefited from the use of antibiotics throughout history. These include traces of tetracycline found in skeletal remains from 350-550 BC, the identification of antibiotic-producing bacteria in the red soils of Jordan (used to treat skin infections) and the anti-malarial qinghaosu (artemisinin) in traditional Chinese medicine used for many ailments (Aminov, 2010; Falkinham *et al.*, 2009). Alexander Fleming discovered the $\beta$-lactam antibiotic penicillin and the first antibiotic to be used in a clinical setting was Prontosil (a sulphonamide) which was produced in 1931 and used to cure a boy dying of staphylococcal septicaemia in 1933 (Gould, 2016). These discoveries marked the beginning of the 'Antibiotic Golden Era' (1940 - 1970) and during World War 2, the War Production Board established large-scale production of penicillin to assist the war efforts (Quinn, 2013). Subsequently there was a boom in the discovery of new antibiotic classes (Figure 1.2), antibiotics were described as wonder drugs and were used widely, with many developed and developing countries recording improved outcomes for infectious disease and, simultaneously, increased life expectancy (Adedeji, 2016).

Antibiotics can act by inhibiting bacterial growth (bacteriostatic), killing the bacterium (bacteriocidal), or by inducing cell lysis (bacterolytic). Figure 1.3, highlights the three key areas of metabolism which antibiotic classes can target; cell wall biosynthesis, nucleic acid synthesis and protein synthesis. Some antibiotics target additional biosynthesis pathways such as mycolic and folic acid synthesis.

Antibiotic resistance is driven by evolution and arises when mutations in DNA result in the bacteria no longer being affected by specific antibiotic compounds, thereby providing an evolutionary advantage for those bacteria that have developed resistance. There are two mechanisms which are the genetic basis for development of resistance; i) random mutations and ii) horizontal gene transfer. Bacteria carrying random mutations which permit survival in the presence of the antibiotic persist and become the predominant species, whilst susceptible bacteria are eliminated

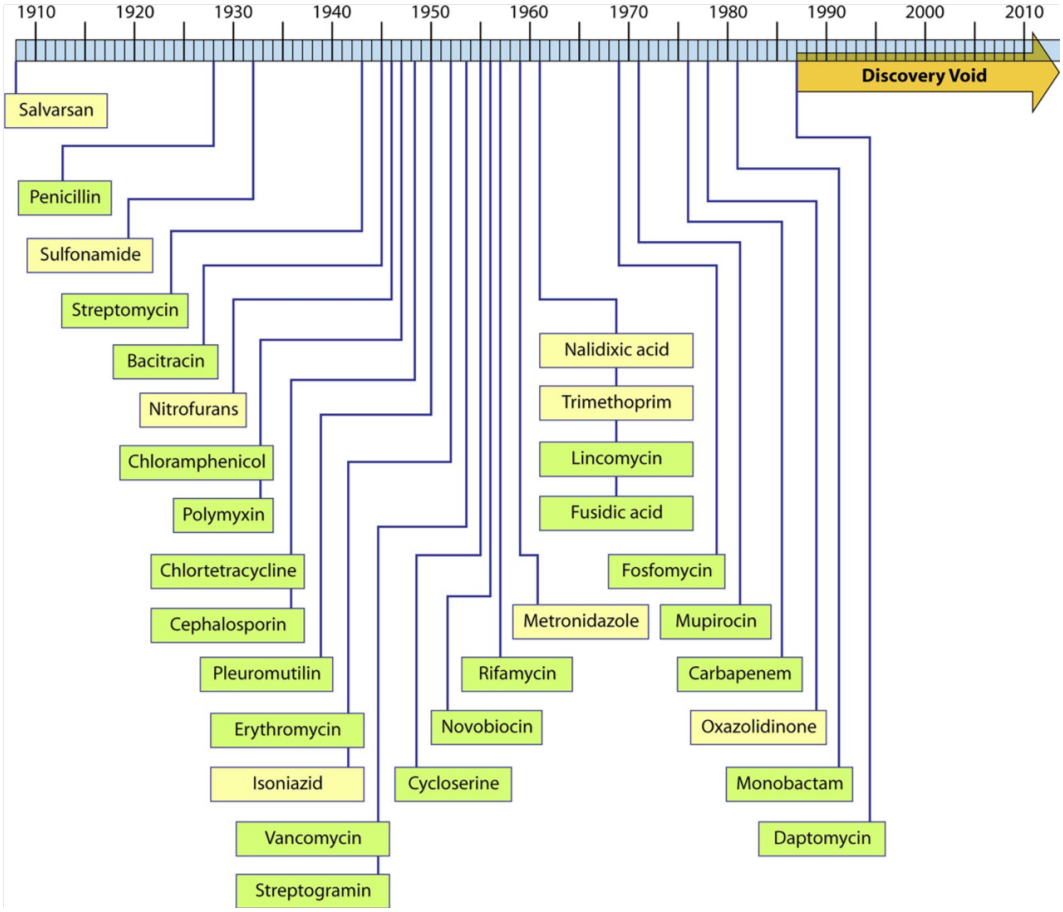



**Figure 1.2: Timeline of antibiotic discovery and the discovery void.** Dates indicated are those of reported initial discovery or patent. Naturally occurring antibiotics are shown in green whilst synthetic antibiotics are shown in yellow. Figure from Silver (2011).

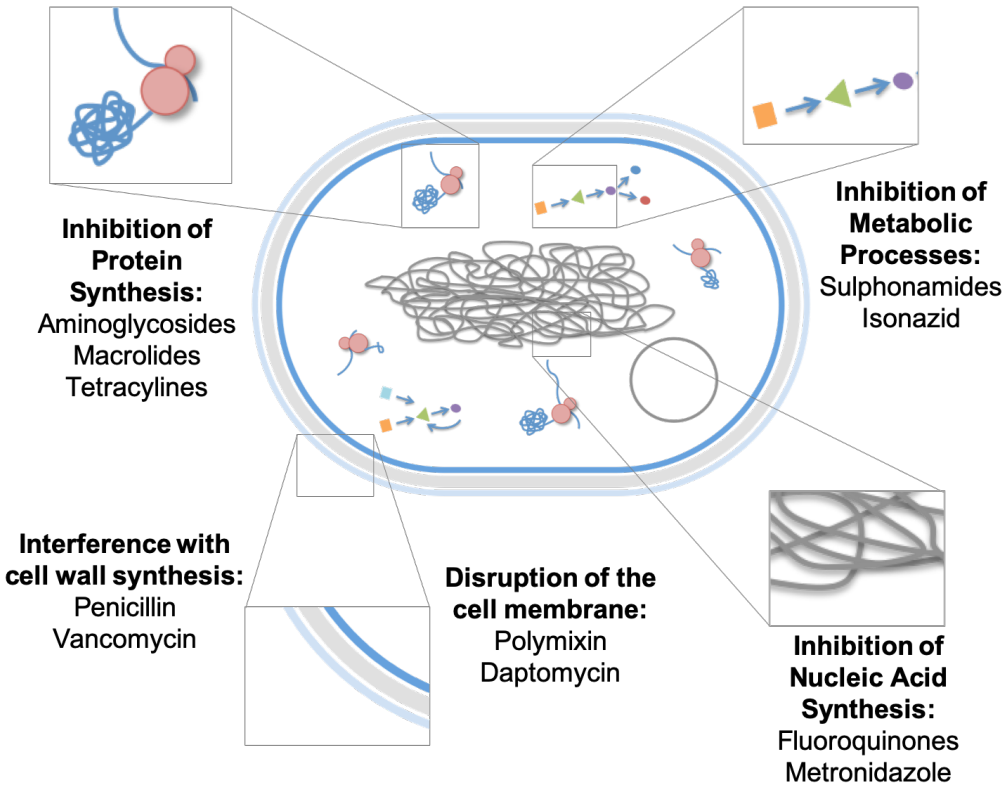



**Figure 1.3: Diagram illustrating the cellular mechanisms which are targets for antibiotics**

by the antibiotic. There is often a fitness cost incurred due to mutations like this and so often they are only maintained in the presence of the antibiotic (Blair *et al.*, 2015). Horizontal gene transfer is where bacteria exchange DNA through transformation, transduction or conjugation. It is one of the most important drivers of bacterial evolution and is a significant cause of antibiotic resistance (Blair *et al.*, 2015). The three key mechanisms which generate bacterial resistance are; preventing access of the antibiotic to the bacterial target, inactivation of the antibiotic and modification of the antibiotic target (Figure 1.1). Preventing access of the antibiotic to its target can occur by either decreasing or replacing non-specific porins, which are responsible for antibiotic uptake, or by increasing efflux pumps, which export antibiotics out of the cell. Enzymes such as $\beta$-lactamases, imipenemase, oxacilinase and carbapenemase hydrolyse the antibiotic thus causing its inactivation, furthermore, aminoglycosides and chloramphenicol are particularly susceptible and can be modified by enzymes such as acetyltransferases, phosphotransferases

and nucleotidyltransferases. Modification of the antibiotic target prevents binding of the antibiotic but without affecting the normal activity of the target. This has been observed in *S. pneumoniae* penicillin binding proteins (PBPs) acquire mutations which prevent the binding of penicillin, generating resistance. Modifications which do not require mutations can also be used to protect the target from antibiotic binding. For more detailed reviews of these mechanisms the author refers the reader to Walsh (2000), Blair *et al.* (2015) and Kapoor *et al.* (2017).

| **Antibiotic** | **Target** | **Mode of action** | **Resistance mechanism** |
|---|---|---|---|
| **Cell wall** | | | |
| $\beta$-Lactams | Transpeptidases/ transglycosylases (PBPs) | Blockade of crosslinking enzymes in peptidoglycan layer of cell walls | $\beta$-lactamases, PBP mutants |
| Vancomycin | D-Ala-D-Ala termini of peptidoglycan and of lipid II | Sequestration of substrate required for crosslinking | Reprogramming of D-Ala-D-Ala to D-Ala-D-Lac or D-Ala-D-Ser |
| **Protein synthesis** | | | |
| Macrolides of the erythromycin class | Peptidyl transferase, centre of the ribosome | Blockade of protein synthesis | rRNA methylation, drug efflux |
| Tetrocyclines | Peptidyl transferase | Blockade of protein synthesis | Drug efflux |
| Aminoglycosides | Peptidyl transferase | Blockade of protein synthesis | Enzymatic modification of drug |
| Oxazolidinones | Peptidyl transferase | Blockade of protein synthesis | Unknown |
| **DNA replication/repair** | | | |
| Fluoroquinolones | DNA gyrase | Blockade of DNA replication | Gyrase mutations to drug resistance |

**Table 1.1: The cellular target and mode of action, and mechanism of bacterial resistance for the main classes of antibiotics.** Figure adapted from Walsh (2000).

The use of antibiotic compounds throughout history has contributed to the accumulation of resistance genes in the human population, and so antibiotic resistance has likely always existed to a greater or lesser extent. Enzymes, known

as $\beta$-lactamases, which hydrolyse $\beta$-lactam antibiotics, originated more than two billion years ago (Aminov, 2010). In his Nobel Prize lecture, Alexander Fleming noted that incorrect use of antibiotics could render them ineffective. This was due to his observation that bacteria could become resistant to antibiotics if exposed to sub-inhibitory concentrations. However, despite this warning, antibiotics have been overused and misused due to inappropriate prescribing in humans, under-dosing through lack of course completion, extensive use in agriculture and possibly by additional factors such as increased use in cleaning products (Ventola, 2015). As a result, antibiotic resistance has risen dramatically since their clinical introduction and a crisis looms. At present there are an estimated 700,000 deaths globally that occur due to resistant bacterial infections; at the current rate of progression this is predicted to rise to 10 million per year by 2050, making antibiotic resistance a global health priority (O'Neill, 2016).

**Antibiotic Resistance in *S. pneumoniae***

Penicillin resistance is a major concern in *S. pneumoniae*, with resistance rates varying between 25-50 % throughout Europe. Interestingly high-level penicillin resistance appears to develop in only a few serotypes; out of the 147 high-level penicillin resistant isolates studied, 90.5 % belong to one of six serotypes (Rauch *et al.*, 1998). $\beta$-lactams such as penicillin covalently bind and inhibit the active site serine of penicillin binding proteins (PBPs). PBPs are involved in the polymerisation and cross-linking of precursors to create the peptidoglycan (PG) layer in the bacterial cell wall. PBP genes have acquired heterologous DNA sequences (some conferring resistance) by horizontal gene transfer, from other bacteria which are indigenous to the upper respiratory tract such as *Streptoccocus mitis* and *Streptococcus oralis* (Ola *et al.*, 2007). This results in mosaic PBPs whose binding site has a reduced affinity for penicillin binding. This is a prime example of antibiotic target modification and shows how the natural competence of *S. pneumoniae* confers an advantage to the bacterium.

Other antibiotics used to treat *S. pneumoniae* infections also show varying levels of resistance. Cephalosporin resistance occurs due to alterations in PBP 2x, 2a and 1a; such resistance increases in parallel with penicillin resistance, with resistance to cefotaxime and ceftriaxone reported for up to 40 % of *S. pneumoniae* strains. Worryingly, resistance to trimethoprim-sulfamethoxazole occurs in 20-60 % of pneumococcal isolates. Whilst resistance to erythromycin and tetracyclines has increased worldwide, prevalence in individual countries is highly variable. Clindamycin resistance is less common than macrolide resistance, but high-level resistance of clindamycin is correlated to macrolide resistance. The *mef(A)* gene, encoding an antibiotic efflux pump, is commonly acquired in these bacteria and leads to macrolide resistance (Reinert, 2009). The original fluoroquinolones had a marginal minimum inhibitory concentration (MIC) against *S. pneumoniae* and were a poor choice of drug for CAP; however, the newer fluoroquinones have improved activity against *S. pneumoniae*, with penicillin resistance having no effect. All strains of *S. pneumoniae* are susceptible to vancomycin and most strains are susceptible to carbapenems, including those resistant to penicillin, making these a viable treatment choice. However, vancomycin tolerance has been observed, and all isolates were also resistant to penicillin (Chenoweth *et al.*, 2000).

Multidrug resistance (MDR) is now endemic in many countries, with 9-25 % of pneumococci in the US found to be multidrug resistant (Chenoweth *et al.*, 2000). It is therefore clear that whilst some treatments for CAP remain effective, resistant strains are increasing at an alarming rate. Therefore there is an urgent need for the development of new antimicrobials.

## 1.3 Peptidoglycan and cell wall biosynthesis

As mentioned previously, cell wall synthesis is one of the main targets for antibiotics. In Gram-positive organisms which lack an outer membrane, the PG is an essential component of the bacterial cell wall which provides strength and rigidity, flexibility

during growth and division, and protects the cell from high osmotic pressures. It is a polymer consisting of alternating carbohydrate residues of *N*-acetyl glucosamine (Glc*N*Ac) and *N*-acetylmuramic acid (Mur*N*Ac) with an appended peptide chain that is cross-linked either directly or indirectly, resulting in the formation of a mesh-like structure (Figure 1.4) (Bugg *et al.*, 2011). PG acts as an anchor for other cell wall components such as surface proteins, CPS and Type I pilus (Dramsi *et al.*, 2008; Larson and Yother, 2017; Shaik *et al.*, 2014).

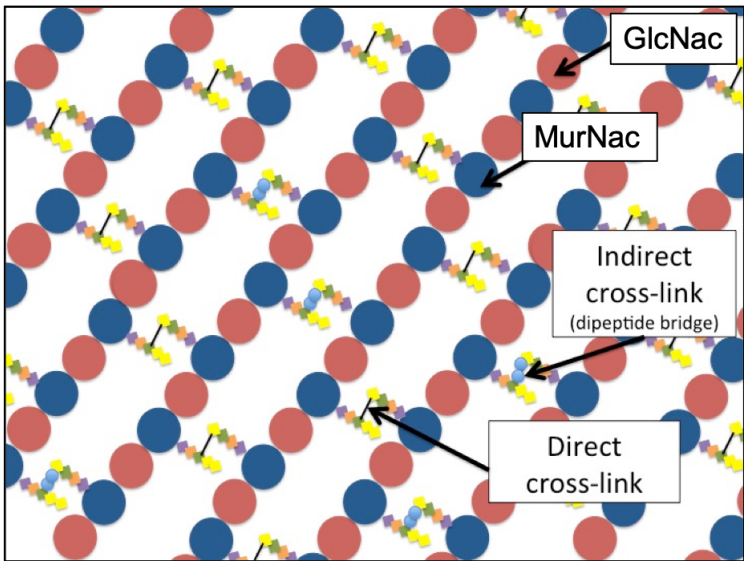



**Figure 1.4: Diagram illustrating the mesh-like structure of peptidoglycan.** Alternating residues of *N*-acetylmuramic acid (Mur*N*Ac) (blue) and *N*-acetylglucosamine (Glc*N*Ac) (red) with direct and indirect cross-links occurring between peptide stems.

Inhibition of cell wall biosynthesis is a key mechanism for many antibiotics, including the most widely used class of antibiotics known as the $\beta$-lactams, as well as glycopeptides and amino acid analogues (Schneider and Sahl, 2010). Figure 1.5 demonstrates the 3 main stages of PG biosynthesis; cytoplasmic, internal face of cytoplasmic membrane and external face of cytoplasmic membrane (Bugg *et al.*, 2011).

Whilst the PG pathway shares similarities in most bacteria, there are minor differences between Gram-positive and Gram-negative bacteria plus additional variations which occur between different species. This section will therefore describe PG biosynthesis broadly, noting variations as necessary with a particular focus on

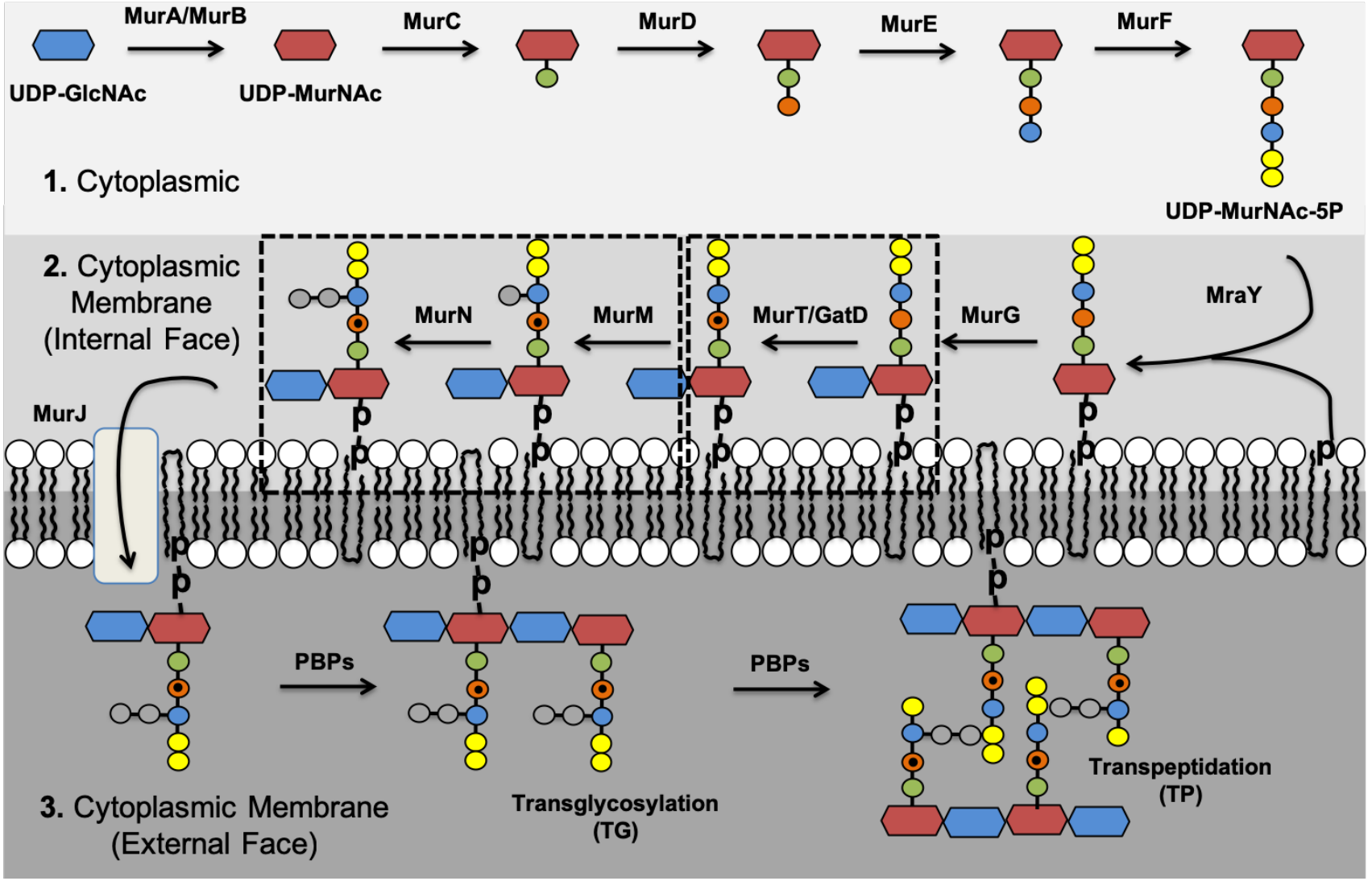



**Figure 1.5: Diagram illustrating stages of the PG Biosynthesis Pathway.** 1) The cytoplasmic stage is characterised by the formation of UDP-*N*-acetylmuramyl pentapeptide (UDP-5P) by the Mur ligases. The pentapeptide (5P) side chain usually comprises L-Ala-$\gamma$D-Glu-L-Lys-D-Ala-D-Ala in Gram-positive organisms. 2) At the internal face of the cytoplasmic membrane MraY catalyses the addition of UDP-5P to undecaprenyl-phosphate forming Lipid I, which is then converted to Lipid II by MurG. The sequence of subsequent reactions, has yet to be elucidated (indicated by the dashed boxes) but, in *S. pneumoniae* the second position $\alpha$ carboxyl of the D-glutamate is amidated to D-iso-glutamine (iGln) by MurT/GatD complex and a dipeptide branch of either L-Ala/L-Ala or L-Ser/L-Ala is appended at the third position lysine by MurM and MurN respectively. Lipid II is translocated across the membrane by MurJ. 3) At the external face of the cytoplasmic membrane PBPs form glycan chains by transglycosylation (TG), with the concomitant release of undecaprenyl-pyrophosphate, and form either direct or indirect cross-links throughout the PG layer via transpeptidation (TP).

those occurring in *S. pneumoniae*.

The cytoplasmic stages of PG synthesis are well characterised, and can be summarised as the production of Park Nucleotide or uridine diphosphate-*N*-acetylmuramyl pentapeptide (UDP-Mur*N*Ac-5P), by formation of UDP-*N*-acetylmuramic acid (UDP-Mur*N*Ac) by MurA and MurB, and the subsequent addition of amino acids by Mur ligases MurC, D, E and F to form a pentapeptide (5P) chain (Barreteau *et al.*, 2008). The pentapeptide chain at this stage of the pathway is most commonly L-Ala-$\gamma$-D-Glu-X-D-Ala-D-Ala, whereby X is L-Lys in Gram-positive bacteria or *meso*-DAP (where DAP denotes diaminopimelic acid) in Gram-negative bacteria. However, other less frequently observed variations of this residue have been documented and are discussed in more detail by Vollmer *et al.* (2008). MraY and MurG subsequently catalyse the conversion of Park Nucleotide into Lipid II(X) (where X can be omitted to refer to all Lipid II molecules, or used to indicate the specific amino acid residue present at the third position of the 5P).

A number of modifications to the 5P of Lipid II can occur, during this stage at the internal face of the cytoplasmic membrane.

The PG of a number of Gram-positive bacteria (including *Mycobacterium tuberculosis*, *S. pneumoniae* and *Clostridium perfringens*) is amidated to contain D-iso-glutamine (iGln) at the second position of the pentapeptide. In *Staphylococcus aureus*, the GatT/MurD enzyme complex was identified to be responsible for the amidation of D-Glu, and homologous transcriptional units have also been identified in other Gram-positive bacteria (Munch *et al.*, 2012; Liu *et al.*, 2017). In *S. pneumoniae* 100 % of the PG contains iGln at the second position of the 5P, which is consistent with the findings that GatT/MurD are essential genes (Liu *et al.*, 2017).

An additional modification which commonly occurs in Gram-positive bacteria is branching, by addition of amino acids to the $\epsilon$-amino group of the third position lysine of the 5P. With the exception of *Weissella viridescens* and *Enterococcus faccium*,

for which branching occurs to the UDP-Mur*N*Ac-5P, this modification occurs to the Lipid II(Lys) (Hegde and Blanchard, 2003; Bellais *et al.*, 2006). The composition and extent of branching varies greatly between species, and as such, there are a number of homologous proteins which are responsible for these amino acid additions (Table 1.2) (Vollmer *et al.*, 2008). The PG of wild type *S. pneumoniae* is heterogeneous, containing branched and unbranched PG. MurM and MurN are the enzymes responsible for the sequential addition of amino acids to generate branched PG precursors (Severin and Tomasz, 1996). MurM aminoacylates the lysine of the 5P with L-Ser or L-Ala and MurN acylates the free amino group of the first amino acid, with an invariable L-Ala. Peptide bond formation occurs in a direction that is reversed compared to peptide bond formation in ribosomal protein synthesis (Filipe *et al.*, 2000a).

| **Enzyme** | **Bridge** | **Species** |
|---|---|---|
| FemX, FemA, FemB | $(Gly)_5$ | *Staphylococcus aureus* |
| BppA1, BppA2 | $(L\text{-}Ala)_2$ | *Enterococcus faecalis* |
| FemX | L-Ala-L-Ser or L-Ala-L-Ser-L-Ala | *Weissella viridescens* |
| MurM, MurN | (L-Ala/L-Ser)-L-Ala | *Streptococcus pneumoniae* |
| $Asl_{fm}$ | D-Asp | *Enterococcus faecium* |
| FemX, FemA, FemB, Lif | $(Gly)_2$-Ser-Gly-Ser | *S. simulans* |
| FemX, FemA, FemB, Epr | $(Gly)_2$-Ser-Gly-Ser | *S. capitis* |
| FemX and VanK | Gly | *S. coelicolor* |

**Table 1.2: Enzymes responsible for branching of peptidoglycan precursors, the nature of the branch and the species to which they belong.** Adapted from (Vollmer *et al.*, 2008; Shepherd and Ibba, 2013b).

The order of amidation and branching reactions has yet to be determined; in *S. pneumoniae* all previous experiments on MurM have utilised non-amidated Lipid II(Lys).

Following synthesis of Lipid II on the internal face of the cytoplasmic membrane it must be transported across the membrane to the external face for incorporation into the PG. There has been much controversy in the field with regards to identification of the flippase, and a number of candidate proteins have been investigated; MurJ,

FtsW and AmJ (Ruiz, 2015). In *Escherichia coli* the flippase has been identified as MurJ, which belongs to the Multidrug/oligosaccharideyl-lipid/polysaccharide exporter family. Structural studies of MurJ show two conformations; inward-open and outward-open which indicate an alternating-access mechanism of transport (Kumar *et al.*, 2019; Kuk *et al.*, 2017; Zheng *et al.*, 2018). YtgP is the MurJ ortholog in Gram-positive bacteria and is essential in *S. pneumoniae* and *S. aureus*. However, *B. subtillis* also encodes the functionally redundant AmJ protein, and so YtgP is not essential (Ruiz, 2015).

Once Lipid II(Lys) is on the external face of the membrane it is incorporated into the PG by the PBPs. PBPs bind penicillin due to the similarity between penicillin and the terminal D-Ala D-Ala of the Lipid II (the PBPs natural substrate) (Figure 1.6).

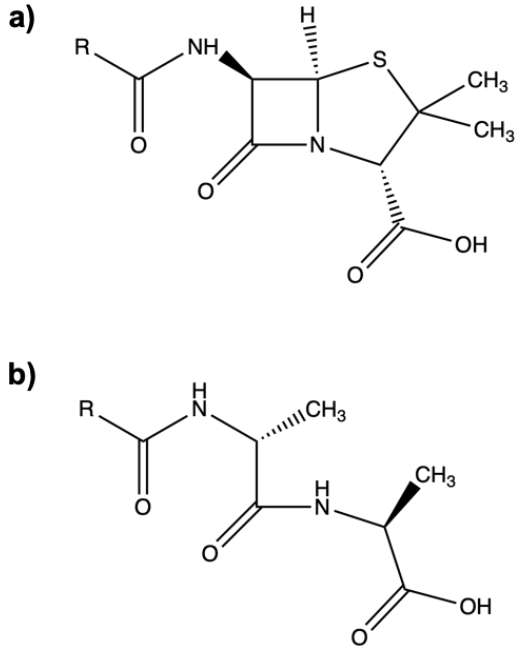



**Figure 1.6: Comparison of Penicillin and D-Ala D-Ala structures.** a) Penicillin and b) D-Ala D-Ala. Figure prepared in ChemDraw (Version 17.1)

The PBPs can be divided into high molecular mass (HMM) and low molecular mass (LMM) PBPs. The HMM PBPs are responsible for PG polymerisation and insertion of PG into an existing wall, and they are divided into Class A and Class B PBPs. Both Class A and B have C-terminal domains with transpeptidase activity.

The N-terminal domain of Class A PBPs is responsible for transglycosylase activity. Class B PBPs interact with cell cycle proteins such as RodA and FtsW which have been shown to possess transglycosylase activity (Meeske *et al.*, 2016; Taguchi *et al.*, 2019). As such the Class A and B PBPs are often referred to as bi-functional and monofunctional respectively. Some bacteria have monofunctional enzymes, with transglycosylase activity, which are similar to the N-terminal domain of Class A PBPs (Sauvage *et al.*, 2008). The LMM PBPs make additional modifications to the PG and are involved in cell separation, peptidoglycan maturation or recycling. LMM PBPs include endopeptidases which separate glycan strands by cleaving cross-bridges, and D,D-carboxypeptidases which act in concert with autolysins to continually remodel and maintain the PG layer (Typas *et al.*, 2011; Sauvage *et al.*, 2008).

Transglycosylation (TG) is the polymerisation of multiple Lipid II precursors to form a glycan chain, whilst transpeptidation (TP) is the formation of cross-bridges between the pentapeptides of individual Lipid II molecules. TP between unbranched or branched Lipid II PG precursors generates direct cross-links or indirect cross-links respectively (Filipe and Tomasz, 2000).

TP between two peptide stems occurs when the peptide bond of D-Ala-D-Ala on the Lipid II (donor) undergoes nucleophilic attack by the active site serine of the PBP forming an acyl-enzyme intermediate. A second Lipid II molecule (acceptor) then attacks the acyl-enzyme intermediate. The nucleophile for this reaction is either the primary amine of lysine (for a direct cross-link) or the amine of the dipeptide branch (for an indirect cross-link). This results in the cross-linking of two Lipid II molecules within the PG (Sauvage *et al.*, 2008).

*S. pneumoniae* contains three bi-functional Class A PBPs; PBP1a, PBP1b and PBP2a. None of these are essential but double mutants (PBP1a and PBP2a) are not viable. Interestingly, whilst the PBPs are able to utilise both Lipid II and iGln Lipid II for TG, the essentiality of MurT/GatD appears to be due to the inability of PBPs to cross-link the peptide stems of Lipid II in the absence of iGln at position 2 (Gisch

*et al.*, 2015).

## 1.4 MurM and branching of peptidoglycan in *S. pneumoniae*

### 1.4.1 Function of MurM

In *S. pneumoniae* the *MurMN* operon encodes two aminoacyl ligases, MurM and MurN, which are responsible for the formation of dipeptide branched Lipid II(Lys) which results in indirectly cross-linked PG. MurM is responsible for the addition of L-Ser or L-Ala to the $\epsilon$-amino group of the third position lysine in the pentapeptide of the Lipid II(Lys). MurN subsequently appends an invariable L-Ala to this amino acid, generating a dipeptide branched Lipid II(Lys) (Fiser *et al.*, 2003; Lloyd *et al.*, 2008). Neither *MurM* or *MurN* are essential; the deletion of *MurN* generates an unusual mono-branched phenotype, whilst deletion of MurM completely abolishes all indirect cross-linking. Whilst MurN is well conserved, MurM is highly mosaic with many different allelic variations between strains. Unlike most bacteria, the PG of *S. pneumoniae* is highly heterogenous and is comprised of both direct and indirect cross-links, in proportions which vary from strain to strain. The key determinant of MurM bias towards incorporation of L-Ala or L-Ser was found to be a 30 amino acid sequence at position 244-274, with particular importance placed on the residue at 260. MurM alleles containing a threonine or lysine at position 260 were found to more commonly incorporate L-Ser or L-Ala into the pneumococcal wall respectively (Filipe *et al.*, 2001b). *In vitro* comparison of MurM alleles from *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* supported this, and confirmed that the amino acid selectivity was not influenced by strain specific differences in the transfer ribonucleic acid (tRNA) pool (Lloyd *et al.*, 2008).

MurM is a $\sim$45 kDa cytoplasmic protein, however in order to access its Lipid II(Lys) substrate it must come into very close proximity with the membrane.

Although the substrate for MurM is Lipid II(Lys) it is important to remember that the hydrophobic lipid tail is embedded in the membrane and the part of the substrate which interacts with MurM is the pentapeptide chain. All previous studies on MurM have been conducted using Lipid II(Lys) so it is unknown whether iGln Lipid II(Lys) is also a substrate. In contrast to the Lipid II(Lys), the second substrate of MurM is a large (Molecular Weight (MW) = 25-30 kDa) hydrophilic tRNA. The different molecular environments and requirements of these two substrates indicate that MurM is a complex and versatile protein. It remains unclear whether MurM carries bound tRNA directly to the membrane or whether it follows a ping-pong mechanism of action, whereby it interacts with the tRNA in the cytoplasm, forming a stable aminoacyl-intermediate which subsequently localises to the membrane to act upon the Lipid II(Lys). MurM has been shown to utilise cognate alanyl-tRNA$^{Ala}$ and seryl-tRNA$^{Ser}$ but interestingly, demonstrated a catalytic preference to seryl-tRNA$^{Ala}$ (Shepherd, 2011). In addition, a minihelix tRNA comprising the acceptor stem and T$\psi$C loop and stem, was enough to support MurM activity (Lloyd *et al.*, 2008). Due to the highly mosaic nature of MurM it is not surprising that the enzyme activity varies between strains; Lloyd *et al.* (2008) demonstrated that aminoacylation of MurM$_{159}$ was $\sim$ 5-fold more active than MurM$_{Pn16}$.

### 1.4.2 Structure of MurM

In *S. aureus*, functional homologues of MurMN such as FemXAB are involved in the synthesis of the pentaglycine cross-bridge. These homologues can provide a useful comparison for both structural modelling and functional elucidation of *S. pneumoniae* MurM. The crystal structure of MurM has not yet been solved, however its predicted structure consists of two domains; the first is a globular domain comprising two twisted $\beta$-sheets surrounded by helices and the second domain is an $\alpha$-helical coiled-coil arm (Figure 1.7). The overall predicted structure shows some similarities to seryl-tRNA synthetases (SerRS) (Fiser *et al.*, 2003) which is

unsurprising, given that tRNA is a common substrate. A crystal structure of *Thermus thermophilus* (*T. thermophilus*) SerRS complexed with tRNA showed that the coiled-coil region made contact with the T$\psi$C loop of the tRNA, directing the acceptor stem into the active site of the synthetase (Biou *et al.*, 1994). The coiled-coil domain of MurM has therefore been suggested to serve an analogous function. However, interestingly, the coiled-coil domain is not essential for tRNA binding in all homologues, since *Weissella viridescens* (*W. viridescens*) FemX does not possess this domain but still utilises tRNA as a substrate (Biarrotte-Sorin *et al.*, 2004). Whilst the second substrate for MurM is Lipid II(Lys), for the *W. viridescens* FemX homologue the second substrate is the soluble UDP-Mur*N*Ac-5P precursor. This suggests that the coiled-coil region in MurM and other FemXAB family members may be important for targeting the protein to the membrane, and hence is not required in *W. viridescens* FemX (Maillard *et al.*, 2005).

**Figure 1.7: Predicted structure of MurM.** MurM is comprised of a globular domain and a coiled-coil domain linked by a flexible unstructured region. Figure adapted from (Fiser *et al.*, 2003).

### 1.4.3 MurM and penicillin resistance

The product of MurM is the substrate for the downstream PBPs, therefore it is unsurprising that MurM has been identified as a non-PBP penicillin resistance determinant (Smith and Klugman, 2001). Whilst the proportion of overall cross-links in the PG varies very little between susceptible and resistant strains of *S. pneumoniae*, the proportion of indirect cross-links (branched PG) correlates strongly with resistance (Garcia-Bustos and Tomasz, 1990). In many penicillin susceptible strains >70 % of the PG contains linear peptides, whilst in contrast, many penicillin resistant strains primarily consist of branched stem peptides (Garcia-Bustos and Tomasz, 1990). However, the presence of highly branched PG does not confer resistance, as demonstrated by *S. pneumoniae* R6 which is a susceptible strain with a high proportion (40 %) of branched peptides in its PG (Garcia-Bustos and Tomasz, 1990). It was proposed that remodelling of the PBP active site, which results in low-affinity PBPs, may also drive a change in substrate preference, promoting the use of hydrophobic branched muropeptides over the more common linear ones (Garcia-Bustos and Tomasz, 1990). Therefore the presence of a functional MurM is important to the development of resistance.

Studies that have investigated the effect of transforming susceptible strains with PBPs and/or MurM from resistant strain donor DNA, all demonstrate that MurM alone is not sufficient for resistance, and that low-affinity PBPs are required. It has been demonstrated that deletion of MurM from penicillin resistant pneumococci results in a dramatically reduced MIC (from resistant to sensitive), and that this resistance can be restored by transformation with plasmid encoded MurM. Additionally, MurM does not show strain specificity, i.e. a MurM from another resistant strain can fully restore the MIC (Filipe *et al.*, 2002). Some studies investigating high-level penicillin resistance (MIC of 16 $\mu g.mL^{-1}$) show that transforming PBPs from a highly resistant strain alone is sufficient to convert a sensitive R6 strain to high level penicillin resistance (Dowson *et al.*, 1995), whilst other studies demonstrate that transforming PBPs alone can only partially restore

resistance, and that MurM is required to achieve full levels of resistance (Smith and Klugman, 2000, 2001).

These conflicting results may have arisen because these experiments were all conducted with different strains, for which both PBPs and MurM are highly mosaic. Categorising MurM alleles as 'altered' MurM and 'standard' MurM, whereby 'altered' can enhance resistance levels in the presence of low-affinity PBPs and 'standard' does not, may be useful. It is possible that the strain that achieved full resistance in the absence of MurM, possessed a 'standard' MurM allele, whilst the strain that demonstrated MurM dependence for restoring high level penicillin resistance, possessed an 'altered' MurM. This may also explain why MurM from R36A (sensitive) can substitute for Pen6 (resistant) MurM (Filipe *et al.*, 2000b). It is possible that, whilst R36A does not contain low-affinity PBPs, and is therefore a sensitive phenotype, it possesses an 'altered' MurM, and so, has the potential to support a resistant phenotype.

Previous studies demonstrated that inactivation of the MurMN operon not only resulted in the disappearance of cross-bridges from pneumococci but also a complete loss of penicillin resistance (Fiser *et al.*, 2003). Taken together, these results suggest that whilst MurM alone is not sufficient for penicillin resistance, the presence of a functional MurM allele is required for the development of penicillin resistance, and an 'altered' MurM allele may provide additional protection against high-levels of penicillin. Reduction in $\beta$-lactam resistance by inactivation of cell wall branching enzymes has also been identified in *S. aureus* (Berger-Bächi and Tschierske, 1998) and *Enterococcus faecalis* (Bouhss *et al.*, 2002). This presents a new paradigm for the development of novel antibiotics, since the inhibition of cell wall branching enzymes such as MurM may restore penicillin sensitivity to previously resistant strains (Fiser *et al.*, 2003).

## 1.5 tRNA and aminoacyl-tRNA synthetases

### 1.5.1 The role of tRNA in protein synthesis and beyond

tRNA is a large macromolecule comprised of 70-96 nucleosides (adenosine (A), guanosine (G), uridine (U) and cytosine (C)), and modified versions thereof (Björk *et al.*, 1987). tRNA is folded into multiple stem-loop structures, whereby nucleotides in the stems are stabilised by Watson-Crick base pairing. The tertiary structure is often represented as a cloverleaf (Figure 1.8a) but actually forms an L-shape in 3D space (Figure 1.8b). Nucleotide positions which are common to all tRNAs are numbered, some of these were identified as invariant and semi-invariant positions, and additional nucleotides can be added in the D-loop or variable loop (Sprinzl *et al.*, 1989). The function of tRNA is to 'carry' amino acids to specific cellular locations for use in metabolic processes. The sequence at the 3' end of all tRNAs is CCA; this is the site where an amino acid is covalently attached to the 2' hydroxyl (OH) or 3'OH of the terminal ribose of the specific tRNA. This reaction is catalysed by the corresponding aminoacyl-tRNA synthetase (aaRS).

During protein synthesis, genetic information encoded on messenger ribonucleic acid (mRNA), in the form of codons, is translated into amino acids by the ribosome. The ribosome matches each mRNA codon with the corresponding anticodon of an aminoacyl-tRNA (aa-tRNA). The ribosome is a large molecular complex made up of ribosomal ribonucleic acid (rRNA) and proteins; upon establishing a codon-anticodon match, the amino acid is transferred from the aa-tRNA to the preceding amino acid generating a polypeptide chain (Figure 1.9). There are many codons in the genome which represent the same amino acid, and the genetic code is therefore referred to as degenerate. These synonymous codons are often used with different frequencies within the genome, which is referred to as codon bias. There are more tRNAs than there are amino acids because some tRNAs exist as isoacceptors (tRNAs possessing alternate anticodons which encode the same amino acid) (Novoa and de Pouplana, 2012). However, whilst there are 61 codons in the universal genetic

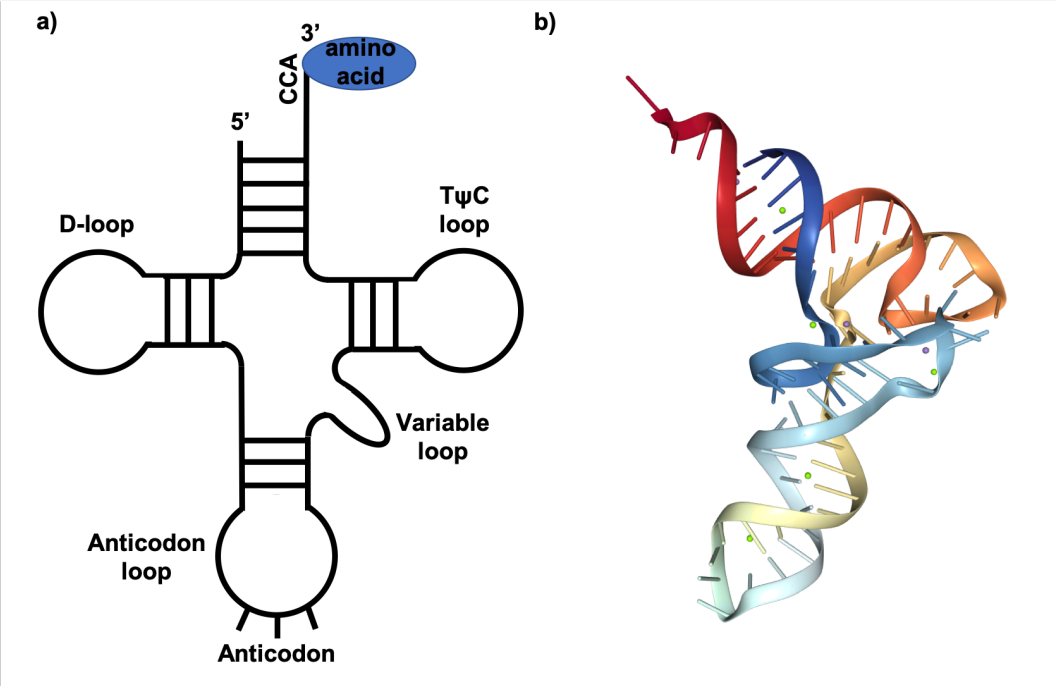



**Figure 1.8: Structure of tRNA** a) The secondary structure of tRNA as represented by a cloverleaf diagram b) The tertiary structure of Yeast phenylalanine tRNA (1EHZ) as determined by X-ray crystallography (Shi and Moore, 2000).

code, there are fewer than 45 different tRNAs in most organisms. Some tRNAs possess anticodons which can recognize multiple, but not necessarily every, codon corresponding to a specific amino acid (Lodish *et al.*, 2008). This phenomonen is explained by the 'wobble hypothesis', for which the third (3') residue of mRNA and the first (5') residue of the anticodon have flexibility in their binding interaction (Crick, 1966).

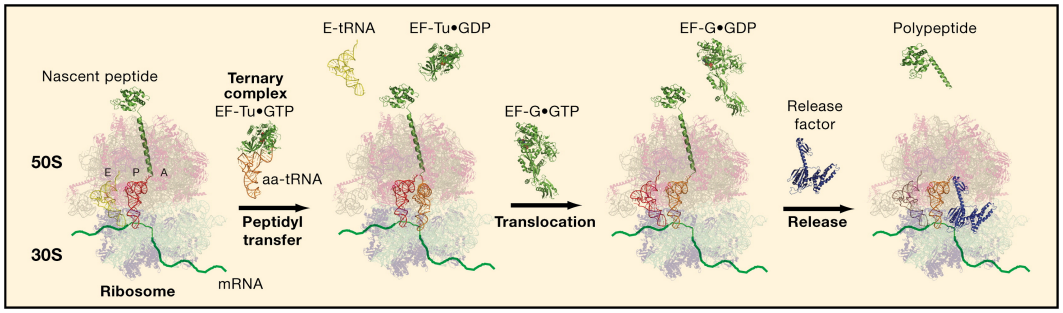



**Figure 1.9: Elongation and termination steps of bacterial translation.** Figure from Zaher and Green (2009).

The fidelity of protein synthesis must be maintained to ensure that proteins fold

and function correctly. Cells have evolved a number of mechanisms to ensure that a high enough fidelity can be maintained whilst minimising the energy expenditure required to do so. Two processes a) codon-anticodon recognition by the ribosome (not discussed here) and b) correct pairing of an amino acid to cognate tRNA by aaRSs, dictate the overall fidelity of protein synthesis, which is estimated to be $10^{-4}$ - $10^{-3}$ (Ibba and Söll, 2000; Bouadloun *et al.*, 1983; Edelmann and Gallant, 1977; Kramer *et al.*, 2010; Kurland, 1992; Yadavalli and Ibba, 2012).

Whilst protein synthesis is the most widely known function for tRNAs, they serve many other non-canonical biological functions. These were classified into 4 main groups; translation, tRNA dependent regulation, genomic interactions and synthetic functions (Katz *et al.*, 2016). Amino acids are provided by aa-tRNAs for a range of non-translational synthetic functions including modification of cellular envelopes, porphyrin biosynthesis, protein degradation pathways, antibiotic synthesis and PG synthesis (Banerjee *et al.*, 2010). Whilst some tRNA isoacceptors can be utilised in both translation and non-canonical functions, some organisms possess specific non-proteogenic isoacceptors. In *S. aureus* unusual $tRNA^{Gly}$ isoacceptors which support PG biosynthesis but not protein synthesis have been identified and demonstrated to bind weakly to elongation factor thermo unstable (EF-Tu) (Bumsted *et al.*, 1968; Giannouli *et al.*, 2009; Levicán *et al.*, 2005). For more information on synthetic or non-synthetic functions of tRNA the author refers the reader to Katz *et al.* (2016).

A significant amount of research has been focused on identifying exactly which elements of tRNA allow aaRSs to discriminate between different tRNAs. These so called 'identity elements' can be positive (determinants) assisting recognition by cognate aaRSs, or negative (anti-determinants) which prevent interactions with non-cognate aaRSs. Identity elements most commonly form direct interactions with aaRSs and are located in the 3' terminus of the acceptor stem and anticodon loop. A total of 40 positions, none of which are found in the continuous acceptor stem or T$\psi$C-loop, have been demonstrated to act as identity elements (Giegé *et al.*, 1998).

Location and extent of identity determinants varies significantly between tRNAs; some are single nucleotide variations such as the discriminator base at position 73, or a unique base pair which is sufficient for recognition (Hou and Schimmel, 1988), others occur as combinations occurring in the anticodon and acceptor stem. In some cases the discriminator base is critical (Pallanck *et al.*, 1992) and in others it is unimportant (Hasegawa *et al.*, 1992). Likewise, whilst most aaRSs require distal interactions with the anticodon for recognition, a few exceptions exist where there are too many isoacceptors with different anticodon nucleotide combinations to provide a unique anticodon:aaRS match (Pang *et al.*, 2014). In addition to direct interactions of identity elements, aaRSs can also recognise specific architectural features belonging to the tRNAs (Giegé *et al.*, 1998).

In addition to the aaRSs, identity elements of tRNAs are also used for recognition by other proteins. For example base-pairing at position 51-63 of tRNA has been shown to modulate affinity to *T. thermophilus* EF-Tu; a protein which facilitates the binding of aa-tRNA to the ribosome (Sanderson and Uhlenbeck, 2007).

### 1.5.2 Aminoacylation by aaRS

The aminoacyl-tRNA synthetases (aaRSs) are proteins responsible for aminoacylating the specific amino acid to its cognate tRNA (also referred to as charging the tRNA). Whilst most aaRS are highly specific for aminoacylation of a single amino acid (aa) to its cognate tRNA, some aaRSs have dual activity and are capable of charging two tRNAs with their cognate amino acids (Lipman *et al.*, 2002). This occurs in a two step process; amino acid activation and transfer (where aa is amino acid):

$$\text{aa} + \text{ATP} \rightarrow \text{aa-AMP} + \text{PP}_\text{i}$$

$$\text{aa-AMP} + \text{tRNA} \rightarrow \text{aa-tRNA} + \text{AMP}$$

In amino acid activation, the carboxyl group of the amino acid attacks the adenosine 5'triphosphate (ATP), forming an aminoacyl-adenylate (aa-AMP) and

releasing inorganic pyrophosphate ($PP_{i)}$. In the second step the aa-AMP undergoes nucleophilic attack by the 2/3'OH on the ribose of the 3' terminal nucleotide of the tRNA (Berg *et al.*, 2002; Li *et al.*, 2015).

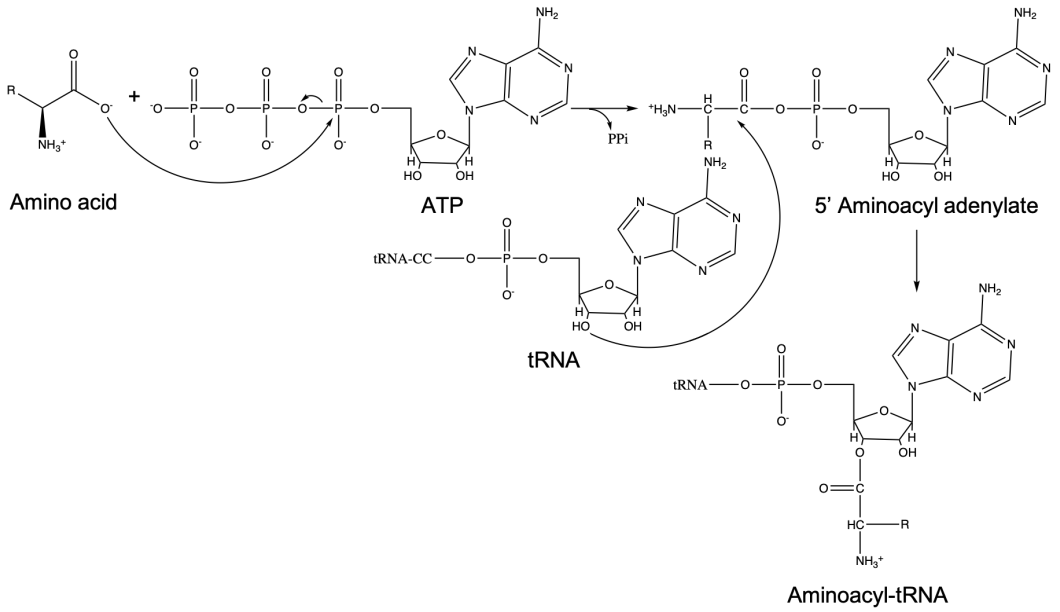



**Figure 1.10: Mechanism of aminoacylation by aminoacyl-tRNA synthetases.** Figure from Li *et al.* (2015)

The aaRS's can be categorised into two major classes based on structural similarities and sequence motifs. The Class I synthetases are mostly monomeric whilst the Class II synthetases are mainly dimeric or multimeric. Class I and II contain 2 and 3 conserved motifs respectively; the motifs are shared within the class but completely absent in the other class (Burbaum *et al.*, 1990; Eriani *et al.*, 1990). Additionally, the OH group on the ribose of the tRNA which catalyses nucleophilic attack on the adenylate is strongly correlated with the Class system, with Class I using the 2'OH and Class II using the 3'OH. There are aaRSs that are exceptions to these classifications, alanyl-tRNA synthetase (AlaRS) and glycyl-tRNA synthetase (GlyRS) only possess one of the three Class II motifs, and whilst phenylalanyl-tRNA synthetase (PheRS) and phosphoseryl-tRNA synthetase (SepRS) possess the Class II motifs, they aminoacylate tRNA at the 2'OH of ribose (Eriani *et al.*, 1990; Englert *et al.*, 2013) (Table 1.3). Despite the Class I and II aaRSs aminoacylating at 2'OH and 3'OH respectively, once appended, the amino acid can migrate back and forth

between the two positions (Schimmel and Söll, 1979). However, interestingly the position of the amino acid can be an important determinant for usage in protein of PG synthesis; the A site of the ribosome is specific for the 3-O-aminoacyl isomer whilst MurM (and homologues from other species) accept only 2′-O-aminoacyl isomers (Fonvielle *et al.*, 2010; Shepherd, 2011).

| **Class I** | **Class II** |
|---|---|
| arginyl-tRNA synthetase (ArgRS) | aspartyl-tRNA synthetase (AspRS) |
| cysteinyl-tRNA synthetase (CysRS) | asparaginyl-tRNA synthetase (AsnRS) |
| glutamyl-tRNA synthetase (GluRS) | histidyl-tRNA synthetase (HisRS) |
| glutaminyl-tRNA synthetase (GlnRS) | lysyl-tRNA (LysRS) |
| isoleucyl-tRNA synthetase (IleRS) | phenylalanyl-tRNA synthetase (PheRS)[b] |
| leucyl-tRNA synthetase (LeuRS) | prolyl-tRNA synthetase (ProRS) |
| methionyl-tRNA synthetase (MetRS) | threonyl-tRNA synthetase (ThrRS) |
| tryptophanyl-tRNA synthetase (TrpRS) | seryl-tRNA synthetase (SerRS) |
| tyrosyl-tRNA synthetases (TyrRS) | alanyl-tRNA synthetase (AlaRS[a]) |
| valyl-tRNA synthetase (ValRS) | glycyl-tRNA synthetase (GlyRS)[a] |
| | O-phosphoseryl-tRNA synthetase (SepRS)[b] |
| | pyrrolysyl-tRNA synthetase (PylRS) |

**Table 1.3: Class I and class II synthetases.** Class I enzymes contain 2 conserved motifs (not found in Class II aaRS's) and aminoacylate at the 2′OH of the terminal nucleotide ribose of the tRNA. Class II enzymes contain 3 conserved motifs (not found in Class I aaRSs) and aminoacylate at the 3′OH of the tRNA's terminal nucleotide ribose. There are however some exceptions (a) indicates synthetases which only have one of the three sequence motifs and (b) indicates Class II enzymes which aminoacyate at the 2′OH of the tRNA's terminal nucleotide ribose. Table modified from Schimmel (1991).

### 1.5.3 Substrate selection by aaRSs

#### tRNA selection by aaRSs

In order to differentiate between cognate and non-cognate tRNAs, aaRSs largely rely on direct interaction with, and recognition of specific identity elements of the tRNA, however recognition of structural features of tRNA have also been observed (Giegé *et al.*, 1998). The coiled-coil region of *T. thermophilus* SerRS was shown to make direct contact with the T$\psi$C loop of the tRNA, directing the acceptor stem into the active site (Biou *et al.*, 1994). Interestingly, successful recognition of the correct

tRNA by the aaRS has also been shown to modulate the affinity for cognate amino acid recognition (Ibba *et al.*, 1996).

**Selecting the cognate amino acid and editing mechanisms**

As mentioned previously correct pairing of an amino acid to its cognate tRNA by the aaRSs is key to maintaining the fidelity of protein synthesis. aaRSs have had to overcome the fact that smaller or isosteric amino acids may also bind in a pocket designed for the cognate amino acid. Examples include, binding of serine and glycine (noncognate) in AlaRS, serine (noncognate) in ThrRS and LysRS, valine (noncognate) in IleRS and threonine (noncognate) in ValRS. Under normal conditions the overall error rate of misaminoacylation by the aaRSs is very low ($\sim 10^{-6}$) which is 100-1000 times more accurate than protein synthesis (Francklyn, 2008). The double sieve mechanism was proposed to explain how aaRS are able to select for their cognate amino acid and deselect for non-cognate amino acids, in order to maintain this high fidelity. Amino acids larger than the substrate were proposed to be excluded at the amino activation site (coarse sieve), but smaller isosteric amino acids may be incorrectly activated. The fine sieve was proposed to be a hydrolytic editing mechanism to remove mis-activated amino acids, whilst retaining the cognate amino acid (Pang *et al.*, 2014). Since some aaRSs are able to activate non-cognate amino acids that are larger, the double sieve mechanism does not provide an accurate representation of this issue. In contrast to a simple sieve mechanism, it is instead a complex interplay of steric, hydrogen bonding and electrostatic forces which regulates the production of correctly acylated tRNA (Francklyn, 2008). For example threonine, serine and valine have very similar molecular weights and side chain stereochemistry; whilst ThrRS commonly mis-aminoacylates with serine, it is able to successfully discriminate between threonine and valine due to the absence of the hydroxyl required for zinc ion coordination on the latter (Berg *et al.*, 2002).

If a non-cognate amino acid is activated *cis*-editing by aaRSs can either occur in

the catalytic site, after activation but before transfer to tRNA (pre-transfer editing) or in a distinct editing site following the transfer of amino acid to the 3′ CCA of the cognate tRNA (post-transfer editing). The requirement for editing varies significantly between different synthetases, and the pre- and post-transfer editing ability of different aaRSs is variable. The rate of amino acid transfer to tRNA has been found to be inversely correlated with the occurrence of pre-transfer editing (Dulic *et al.*, 2010). Additionally, the rate of translocation to the editing site does not vary between tRNA mis-acylated with different amino acids (Nomanbhoy and Schimmel, 2000).

AlaRS and ThrRS mis-acylate serine and so rely on editing activities to maintain the fidelity of protein synthesis. Mis-aminoacylation with serine or glycine by AlaRS occurs at an error rate of 1/500 or 1/250 respectively. As a result many aaRSs have a distinct editing domain which can correct mis-acylation. The editing domain of AlaRS weakly aligns with that of ThrRS; in AlaRS, but not ThrRS, the editing domain is found in all kingdoms of life, and so may have been acquired by ThrRS during evolution (Beebe *et al.*, 2003). Cys666 and Cys182 have been identified as critical for hydrolytic editing in AlaRS and ThrRS respectively (Beebe *et al.*, 2003; Ling and Söll, 2010; Wu *et al.*, 2014). Cys443 was also identified as necessary for editing of alanine in ProRS, and additional sites (His73, His77 and His186) proposed to be important in zinc coordination, are important for editing in ThrRS (Beuning and Musier-Forsyth, 2000; Dock-Bregeon *et al.*, 2000).

Additional free-standing editing domain homologues have been identified for some aaRSs (An and Musier-Forsyth, 2005; Ahel *et al.*, 2003). Whilst the AlaXs are the most widespread *trans*-editing factors, interestingly, *S. pneumoniae* does not encode any known AlaXs. It is therefore likely that another mechanism exists to help maintain the fidelity of protein synthesis in this bacteria (Shepherd and Ibba, 2013a).

Despite numerous quality control mechanisms, decreased translational fidelity can arise as a result of a genetic mutation, aminoglycoside antibiotics, and a variety of environmental stresses such as nutrient deprivation and oxidative stress (Mohler

and Ibba, 2017). Treatment of ThrRS with $H_2O_2$ has been shown to abolish editing activity, resulting in an increase in seryl-tRNA$^{Thr}$ (Wu *et al.*, 2014).

## 1.6 Oxidative stress in *S. pneumoniae*

Obligate anaerobes and microaerophiles cannot grow in the presence of oxygen ($O_2$), in addition, the growth of aerobic bacteria may be severely affected during hypoxia. This demonstrates that molecular ($O_2$) is highly toxic; whilst ($O_2$) itself does not react with most bio-molecules, toxicity arises from its partial reduction, by the progressive addition of electrons ($e^-$), into ($O_2$) species that are more reactive than itself (Figure 1.11) (Imlay, 2003; Mishra and Imlay, 2012). Almost all organisms possess enzymes that scavenge superoxide ($O_2^-$) and $H_2O_2$ to provide protection from oxidative stress (Mishra and Imlay, 2012). Oxidative stress can come from a range of exogenous sources including $H_2O_2$-secreting lactic acid bacteria, NADPH oxidase responses of plants and macrophages, and redox reactions from phytochemistry (Mishra and Imlay, 2012).

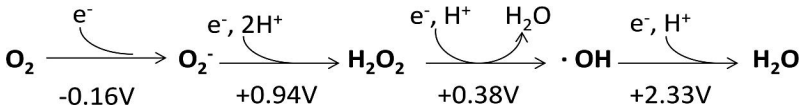



**Figure 1.11: The reduction of oxygen to form reactive oxygen species (ROS).** Oxygen is reduced by sequential addition of electrons to form superoxide, hydrogen peroxide, hydroxyl radical and finally water. Figure from Mishra and Imlay (2012).

Under aerobic conditions, bacteria convert glucose into lactic acid with the concomitant generation of ATP. Lactate oxidase (Lox) converts lactic acid to pyruvate, generating $H_2O_2$ as a metabolic biproduct (Seki *et al.*, 2004). Pyruvate can subsequently be converted to acetate via the thioclastic pathway, or by pyruvate oxidase generating additional ATP. NADH oxidase (NOXase) directly catalyzes the four-electron reduction of $O_2$ to water ($H_2O$) and is important for aerobic growth (Gibson *et al.*, 2000). In many lactic acid bacteria, Lox NOXase and pyruvate oxidase are only active when the glucose source has been exhausted (Taniai and Yoshida, 2008).

$H_2O_2$ can also be converted into a highly reactive, and dangerous, hydroxyl radical ($HO^{\bullet}$ by the Fenton-reaction:

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + HO^{\bullet} + OH^{-}$$

$$Fe^{3+} + H_2O_2 \rightarrow Fe^{2+} + HOO^{\bullet} + H^{+}$$

$O_2^-$ is also formed during aerobic growth by autoxidation of the respiratory dehydrogenases. During normal turnover of dehydrogenases, if $O_2$ collides with a reduced flavin before it has transferred the electron to the next carrier, the $e^-$ can partially reduce oxygen. The concentration of $O_2$ will therefore dictate how quickly $O_2^-$ is produced (Imlay, 2003).

Reactive oxygen species (ROS) target many biological molecules including DNA, RNA, proteins and lipids resulting in high mutation rates, growth defects and cell death (Imlay, 2003; Mishra and Imlay, 2012).

Due to high prevalence of ROS, there are a number of major response regulators that provide mechanisms to protect against oxidative stress. OxyR, SoxRS, and RpoS in *E. coli* have been shown to be functionally conserved in proteobacteria to Actinobacteria (Chiang and Schellhorn, 2012). Antioxidant enzymes such as superoxide dismutases (SOD), catalase and hydroperoxidases, as well as DNA repair mechanisms, provide protection against ROS in many bacteria. However, if the oxidative stress exceeds the capacity of the bacterial response then the effects will be toxic. In *E. coli*, low concentrations of $H_2O_2$ activate OxyR, resulting in the expression of catalase, a protein that converts $H_2O_2$ to $H_2O$ and $O_2$, and transcriptional chaperones such as DnaK, which increase tolerance to aberrant proteins.

However, despite these responses, under severe oxidative stress, bimodal killing of *E. coli* is observed by $H_2O_2$ (Ling and Söll, 2010). 0.2-0.5 mM $H_2O_2$ can induce protein mistranslation, and in *E. coli* $H_2O_2$ has been shown to cause oxidation of Cys182 of the ThrRS, impairing its editing capacity and increasing mis-aminoacylation of $tRNA^{Thr}$ with serine (Ling and Söll, 2010).

*S. pneumoniae* is unique in that its Lox and Streptococcus pyruvate oxidase (SpxB) genes are constitutively expressed, resulting in generation of more ATP but at the expense of producing copious amounts of $H_2O_2$. Figure 1.12 shows the proposed pathway of aerobic metabolism in *S. pneumoniae* (Taniai and Yoshida, 2008).

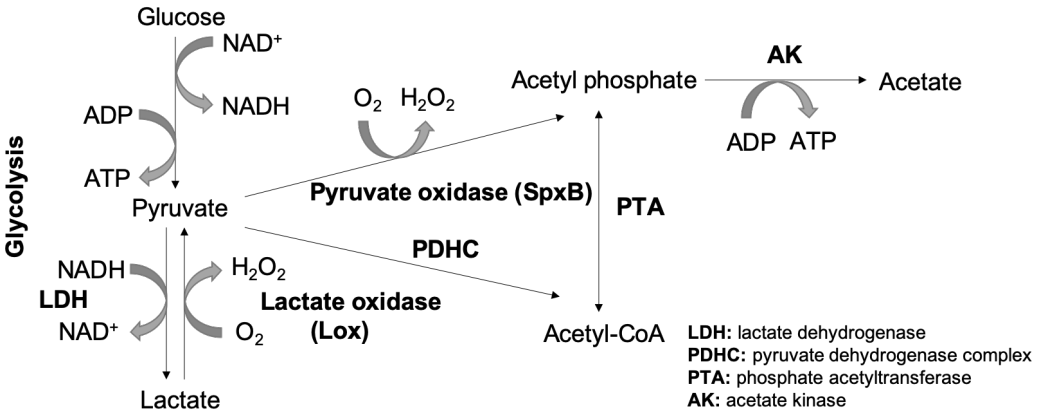



**Figure 1.12: Proposed pathway of aerobic metabolism of glucose in *S. pneumoniae*.** Figure from Taniai and Yoshida (2008)

*S. pneumoniae* produces high levels (∼1.0 mM) of exogenous $H_2O_2$, killing or inhibiting growth of other organisms such as *S. aureus*, *Haemophilus influenza*, *Morexella catarrahais* and *Neisseria meningitidis* that are inhabiting the same microenvironment (Pericone *et al.*, 2000; Sel, 2009). It is intriguing therefore that *S. pneumoniae* does not possess the global regulators (OxyR or PerR), or either of the $H_2O_2$ scavenging enzymes (catalase and NADH peroxidase) that are present in other species (Pericone *et al.*, 2003). The mechanism by which *S. pneumoniae* survives in high exogenous levels of $H_2O_2$ remains largely unknown, however pyruvate oxidase (SpxB) is required for resistance to exogenous $H_2O_2$ and an aquaporin (So-aqpA) is important for $H_2O_2$ detoxification (Pericone *et al.*, 2003; Tong *et al.*, 2019).

## 1.7 Hypothesis and Aims

High levels of $H_2O_2$ stress experienced by *S. pneumoniae*, along with few protective mechanisms against oxidative stress have been characterised. *S. pneumoniae* has been shown to experience high levels of $H_2O_2$ stress, and very few protective mechanisms against oxidative stress have been characterised. In addition $H_2O_2$ has been shown to increase mis-aminoacylation and translational error in other organisms. Therefore we propose that MurM shows a preference for mischarged tRNAs, such as seryl-tRNA$^{Ala}$ and seryl-tRNA$^{Thr}$ as a mechanism for maintaining translational fidelity during $H_2O_2$ stress, by directing mischarged tRNAs away from protein synthesis into PG synthesis (Figure 1.13).

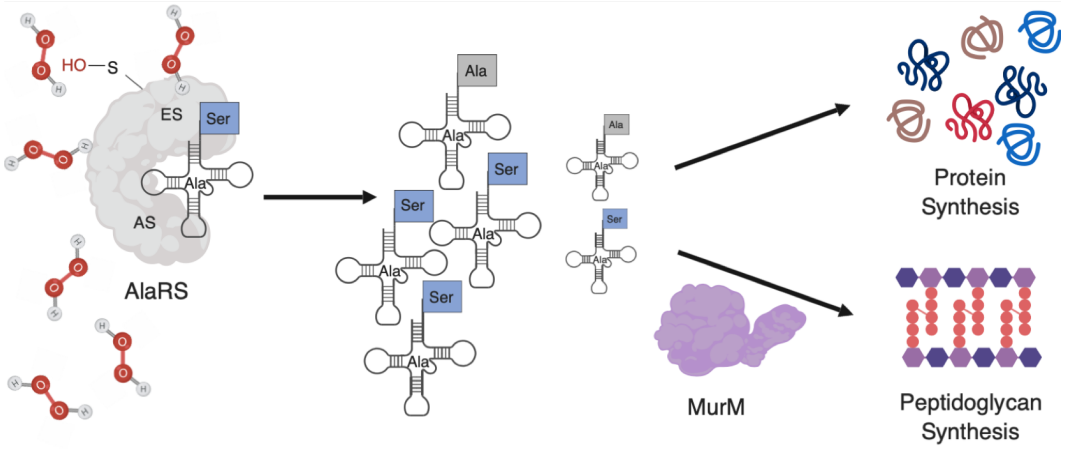



**Figure 1.13: Proposed hypothesis of how MurM helps maintain the fidelity of protein synthesis in the presence of oxidative stress.** The presence of hydrogen peroxide ($H_2O_2$) results in the oxidation of the cysteine residue in the editing domain of the aaRS. This causes an increase in mis-aminoacylated tRNA species which are eliminated by incorporation into the peptidoglycan by MurM. This diagram uses AlaRS and mis-aminoacylation with serine as an example. AS = active site, ES = editing site. Figure created with BioRender.com

The main aims were to characterise the effect of $H_2O_2$ on AlaRS and ThrRS, to investigate substrate specificity of MurM for cognate versus non-cognate aa-tRNAs and Lipid II(Lys) versus iGln Lipid II(Lys). Computational modelling aimed to provide a better understand of the structure of MurM and any possible sites of interaction with substrate, the membrane and other proteins.

# Chapter 2

# Materials and methods

The author would like to acknowledge the help of the Media Preparation Facility in the School of Life Sciences at the University of Warwick, for providing media, and standard buffers or solutions.

This chapter will outline the final, fully optimised protocols used to conduct the research reported here. Where optimisation was required, the experimental procedures and any interim results will be detailed and discussed in the appropriate results chapter.

Many protocols and methods remain constant despite their broad range of application. These standard methods will be outlined initially and any adjustments or modifications to such methods will be discussed subsequently in direct reference to the specific experiment.

All cell culture manipulations were conducted using appropriate aseptic techniques inside a biological safety cabinet (Richmond Scientific).

## 2.1 Chemicals, reagents, buffers and growth media

### 2.1.1 Chemicals and reagents

Unless otherwise stated all chemicals were supplied by Sigma-Aldrich (UK), Fisher Scientific (UK) or Melford Laboratories Ltd.

### 2.1.2 Buffers

Buffers were prepared using MilliQ $H_2O$ and pH was adjusted using a SevenEasy pH meter (Mettler Toledo) calibrated with pH 4.01, 7.01 and 10.01 buffer solution standards (Thermo Scientific). Buffers used for protein purification were filtered using a 0.2 $\mu$m MF-Millipore Membrane Filter, and buffers used for tRNA preparations were autoclaved at 15 psi for 20 minutes (mins) at 120 °C prior to use.

### 2.1.3 Growth media

Unless stated otherwise, where required antibiotics were used at a final concentration of 100 $\mu$g.mL$^{-1}$, 50 $\mu$g.mL$^{-1}$ and 35 $\mu$g.mL$^{-1}$ for Ampicillin (Amp), Kanamycin (Kan) and Chloramphenicol (Cam) respectively in both solid and liquid media cultures.

#### Liquid culture media

All liquid culture media used throughout this project were prepared as detailed below:

- **Luria-Bertani (LB) media (Bertani, 1951):** 1.25 % (w/v) Bacto-tryptone, 0.625 % (w/v) yeast extract and 1.25 % (w/v) sodium chloride (NaCl) at pH 7.5.

- **Super Optimal Broth with Catabolite repression (SOC) media:** 2 % (w/v) Bacto-tryptone, 0.5 % (w/v) Bacto-yeast extract, 0.05 % (w/v) NaCl, 2.5 mM potassium chloride (KCl) and 20 mM glucose at pH 7.0.
- **Tryptic Soy Broth (TSB) media:** 17 $g.L^{-1}$ casein peptone (pancreatic), 2.5 $g.L^{-1}$ dipotassium hydrogen phosphate ($K_2HPO_4$), 2.5 $g.L^{-1}$ glucose, 5 $g.L^{-1}$ NaCl and 3 $g.L^{-1}$ soya peptone.
- **Brain Heart Infusion Media (BHI):** 37 $g.L^{-1}$ of BHI broth.
- **Auto Induction Medium (AIM) 2YT Broth base including trace elements:** 4.6 % (w/v) of dehydrated AIM medium (FORMEDIUM$^{TM}$).

All media was sterilised by autoclaving at 120 °C for 15 mins.

#### Solid growth media

All solid media used throughout this project were prepared as detailed below:

- **LB agar (LBA):** LB medium with 1.5 % (w/v) bacto-agar.
- **Muller-Hinton horse blood agar:** 38 $g.L^{-1}$ Muller-Hinton agar (Merck) with 5 % (v/v) defibrinated horse blood (EO Laboratories Limited).
- **TSB agar:** TSB media with 1.5 % (w/v) bacto-agar.
- **BHI agar:** BHI media with 1.5 % (w/v) bacto-agar.

Agar was sterilised by autoclaving at 120 °C for 15 mins. Agar was cooled to <50 °C, appropriate antibiotics were added and plates were poured and left to dry in a biological safety cabinet.

## 2.2 Characterising *S. pneumoniae* strains

All antimicrobial susceptibility testing was conducted in collaboration with John Moat at the Antimicrobial Screening Facility, University of Warwick.

### 2.2.1 Susceptibility testing by disk diffusion method

Glycerol stocks of *S. pneumoniae* (159) and *S. pneumoniae* (Pn16) were single colony streaked onto Muller-Hinton horse blood agar plates and a single Oxacillin disk (Oxoid™) was placed on top of the area with the highest microbial load using tweezers. Plates were incubated at 37 °C, overnight with 5 % (v/v) carbon dioxide ($CO_2$). The zone of inhibition surrounding the disk was measured and the susceptibility determined in accordance with the European Committee on Antimicrobial Susceptibility Testing (EUCAST) disk diffusion method (EUCAST).

### 2.2.2 Minimum inhibitory concentration determination

#### Preparation of inoculum

Glycerol stocks of *S. pneumoniae* (159) and *S. pneumoniae* (Pn16) were plated onto Muller-Hinton horse blood agar plates and incubated at 37 °C with 5 % (v/v) $CO_2$ for 18 hours. ATCC *S. aureus* 29213 (for use as a positive control) was plated onto LBA plates and incubated at 37 °C overnight. Cells were harvested using a cotton swap and resuspended in phosphate-buffered saline (PBS) to a turbidity matched to a McFarland standard 1.

#### Minimum inhibitory concentration determination by E-strip method

Inoculum was plated onto Muller-Hinton horse blood agar plates and a Penicillin G (PenG) E-strip (Thermo Scientific™ M.I.C Evaluator™) was applied to the centre of the plate using tweezers. As a positive control a PenG E-strip was applied to ATCC *Staphylococcus aureus* 29213. Plates were photographed and the zones of inhibition were interpreted in accordance with Clinical Laboratory Standards Institute (CLSI) (CLSI).

**Minimum inhibitory concentration determination by broth microdilution method**

A standard broth microdilution method was used in accordance with CLSI protocol (CLSI). MIC values were determined using standard methods and data was interpreted in accordance with the EUCAST breakpoints (EUCAST). ATCC *Staphylococcus aureus* 29213 was used as a positive control. Strains were classified as resistant or sensitive in accordance with EUCAST (EUCAST).

### 2.2.3 Whole genome sequencing of *S. pneumoniae* strains

The work outlined in this section was conducted in collaboration with Catherine Rowland (University of Warwick). All sequencing was conducted at Micropathology Ltd with the guidance and supervision of Dr Daniel Hand, and the bioinformatics analysis was conducted by Dr Daniel Hand (Micropathology Ltd.)

#### *S. pneumoniae* growth and DNA extraction

10 mL BHI medium with 5 % (v/v) defibrinated Horse Blood (EO Laboratories Limited) was innoculated with 500 $\mu$L glycerol stocks of *S. pneumoniae* (159) and *S. pneumoniae* (Pn16). Cultures were incubated at 37 °C overnight, with 5 % (v/v) $CO_2$. Genomic DNA was prepared using Maxwell® 16 DNA Purification Kits in accordance with the manufacturers protocol.

#### Next generation sequencing

Deoxyribonucleic acid (DNA) was size and purity analysed using Fragment Analyzer (Advanced Analytical) and quantified using an Invitrogen™Qubi™ 3.0 Fluorometer. Library preparation was conducted in accordance with Nextera® XT DNA Library Prep Kit Reference Guide. Next generation sequencing (NGS) was

conducted using the Nextera® XT MiSeq® Reagent Kit Version 3 on the Illumina MiSeq™ instrument.

***De novo* genome assembly**

Genomes were *de novo* assembled using SPAdes (Version 3.11.1) (Bankevich *et al.*, 2012), assessed using QUAST (Version 4.1) (Gurevich *et al.*, 2013) and annotated with Prokka (Version 1.12) (Seemann, 2014).

The *de novo* assembled sequences of MurM from *S. pneumoniae* (159) and *S. pneumoniae* (Pn16) were aligned with the previously published sequences (Lloyd *et al.*, 2008) using CLUSTAL OMEGA (Version 1.2.4) multiple sequence alignment tool (Sievers *et al.*, 2011).

## 2.3 Manipulation of DNA

All construct details are presented in Table 2.1. All primers were synthesised by Integrated DNA Technologies (IDT) and details can be found in Appendix 7.1. All polymerase chain reactions (PCRs) were conducted using a SureCycler 8800 thermocycler (Agilent Technologies). Annealing temperatures for PCR conditions were calculated using New England Biolabs (NEB) Tm Calculator (Version 1.9.13) (https://tmcalculator.neb.com).

### 2.3.1 DNA sequencing of plasmid constructs

All sequencing was conducted by GATC Biotech (Eurofins) with 80 - 100 ng.µl$^{-1}$ plasmid or 20 - 80 ng.µl$^{-1}$ PCR product and 5 pmol.µl$^{-1}$ primer. The quality of sequence reads was analysed and assembled manually using A plasmid Editor (ApE) (Version 2.0.55). The final sequence was translated to the corresponding amino acid code using ExPASy Translate Tool (Gasteiger E., 2003) and aligned to

the construct reference sequence using CLUSTAL OMEGA (Version 1.2.4) multiple sequence alignment tool (Sievers *et al.*, 2011).

| **Construct Name** | **Description** | **Antibiotic selection** | **Affinity Tag** |
|---|---|---|---|
| pET28a::MurA | *E. coli* MurA in pET28 vector | Kan | C-terminal 6-His |
| pET26a::MurB | *P. aeruginosa* 01293 MurB in pET26a vector (El Zoeiby *et al.*, 2001) | Kan | C-terminal 6-His |
| pET30a::MurC | *P. aeruginosa*01 MurC in pET30a vector (El Zoeiby *et al.*, 2000) | Kan | C-terminal 6-His |
| pET21b::MurD | *P. aeruginosa*01293 MurD in pET21b (El Zoeiby *et al.*, 2001) | Amp | C-terminal 6-His |
| pET21b::MurE | *S. pneumoniae* (Pn16) MurE in pET21b (A.J. Lloyd, Warwick) | Amp | C-terminal 6-His |
| pPROEX::MurF | *S. pneumoniae* (Pn16) MurF in pPROEX-HTa vector (Lloyd *et al.*, 2008) | Amp | N-terminal 6-His |
| pET26b::AlaRS | *S. pneumoniae* (159) AlaRS in pET26b vector (Lloyd *et al.*, 2008) | Kan | C-terminal 6-His |
| pET26a::AlaRS (mutant) | *S. pneumoniae* (159) AlaRS in pET26a vector containing mutations Cys669Ala and His587Gln | Kan | C-terminal 6-His |
| pET21b::SerRS | *S. pneumoniae* (Pn16) SerRS in pET21b vector (Lloyd *et al.*, 2008) | Amp | C-terminal 6-His |
| pET22b::ThrRS (spr1472) | *S. pneumoniae* ATCC BAA-255/R6 ThrRS in pET22 vector | Amp | C-terminal 6-His |
| pET22b::gMurM (159) | gBlock of *S. pneumoniae* (159) MurM sequence (Lloyd *et al.* (2008)) in pET22b vector | Kan | TEV cleavable C-terminal 6-His |
| pET22b::gMurM (Pn16) | gBlock of *S. pneumoniae* (Pn16) MurM sequence (as published in Lloyd *et al.* (2008)) in pET22b vector | Kan | TEV cleavable C-terminal 6-His |
| pET22b::gMurM (Pn16)* | gBlock of *S. pneumoniae* (Pn16) MurM containing mutations Met134Lys and Arg135Gln | Kan | TEV cleavable C-terminal 6-His |

**Table 2.1: Gene constructs used in this project.** All constructs will be referred to as detailed in the construct name.

### 2.3.2 Colony PCR

Single colonies were re-streaked on LBA plates containing the appropriate antibiotic selection and incubated at 37 °C overnight. Half of the re-streaked colonies were harvested into 50 $\mu$L $H_2O$ and incubated at 100 °C for 10 mins. PCR mixes comprised 1 x Standard *Taq* Reaction Buffer (NEB), 1.5 mM magnesium chloride ($MgCl_2$), 0.5 mM dNTPs, 0.5 $\mu$M Forward Primer, 0.5 $\mu$M Reverse Primer, 1.25 units (U) Taq Polymerase, and 5 $\mu$L Template DNA (boiled colony mix). Final PCR conditions can be found in Table 2.2.

| **Component** | **Temperature (°C)** | **Time (min:sec)** |
|---|---|---|
| Initial denaturation (Hot Start) | 95 | 05:00 |
| Start cycle (x 30) | | |
| Denaturation | 95 | 01:00 |
| Annealing | 40 | 01:30 |
| Extension | 72 | 03:00 |
| End cycle | | |
| Final Extension | 72 | 05:00 |
| Hold | 4 | ∞ |

**Table 2.2: Colony PCR conditions**

### 2.3.3 Visualisation of DNA

0.8 - 2 % (w/v) electrophoresis grade agarose (Invitrogen) was suspended in 40 mM Tris, 20 mM acetic acid, and 1 mM ethylenediaminetetraacetic acid (EDTA) (pH 8.3) and solubilised by microwaving. When cooled to ∼50 °C, 1 x GelRed® nucleic acid stain was added to the mixture which was cast as a flat bed gel and left to set at 4 °C.

Samples were prepared using 1 x NEB Gel Loading Dye (no sodium dodecyl sulfate (SDS)) and an appropriate ladder (Quick-Load® 1 kb DNA Ladder, Quick-Load®

100 bp DNA Ladder (NEB)) prepared in accordance with the manufacturers protocol.

Gels were assembled in a MultiSUB Midi Horizontal Gel System (Cleaver Scientific). Samples were loaded to each well and the gels were run at 100 volts (V) for 60 mins in 40 mM Tris, 20 mM acetic acid, and 1 mM EDTA (pH 8.3). Gels were visualised using Syngene G:Box and GeneSnap software (Version 7.12).

### 2.3.4 Gel extraction of DNA

Gels were visualised using a UVP Benchtop UV Transilluminator, bands of interest were cut from the gel using a scalpel and extracted using Monarch® DNA Gel Extraction Kit as per the manufacturers protocol.

### 2.3.5 Purification of PCR products

PCR products were purified using Monarch® PCR and DNA Cleanup Kit as per the manufacturers protocol.

### 2.3.6 Quantification of DNA

1 $\mu$L sample of DNA was quantified using IMPLEN NanoPhotometer®.

### 2.3.7 Plasmid purification

A single colony, from a transformation plate (Section 2.4.2) was used to inoculate 5 mL LB media, containing appropriate antibiotic selection, and incubated at 37 °C, shaking at 180 revolutions per minute (rpm), overnight. Cells were pelleted by centrifugation at 1,800 *xg* for 10 mins. Plasmid was isolated from cell pellets using Monarch® Plasmid Miniprep Kit in accordance with the manufacturers protocol.

Plasmid DNA was isolated from 5 mL LB media cultures using Monarch® Plasmid Miniprep Kit as per the manufacturers protocol.

### 2.3.8 Gibson cloning of proteins

$MurM_{Pn16}$ and $MurM_{159}$ gBlocks® gene fragments containing a 5′ NdeI and 3′ XhoI restriction sites, a C-terminal Tobacco Etch Virus (TEV) protease cleavage site and a C-terminal hexa-histidine (6-His) tag were designed, codon optimised using the GenScript Rare Codon Analysis Tool (https://www.genscript.com/tools/rare-codon-analysis) and synthesised by IDT (Appendix 7.1). PCR mixes comprising 1 x Q5 Master Mix, 0.5 $\mu$M Forward Primer, 0.5 $\mu$M Reverse Primer, 20 ng template DNA and 3 % (v/v) dimethyl sulfoxide (DMSO) were used to amplify gBlocks® and pET22b. Final PCR conditions can be found in Table 2.3 and annealing temperatures of $MurM_{Pn16}$ and $MurM_{159}$ gBlocks®, and pET22b amplifications were 65.5 °C, 64.5 °C and 64.5 °C respectively. PCR products were purified and quantified as per Section 2.3.5 and 2.3.6.

| **Component** | **Temperature (°C)** | **Time (min:sec)** |
|---|---|---|
| Initial Denaturation | 98 | 00:10 |
| Start Cycle (x 30) | | |
| Denaturation | 98 | 00:10 |
| Annealing | variable | 00:30 |
| Extension | 72 | 00:30 |
| End Cycle | | |
| Final Extension | 72 | 03:00 |
| Hold | 4 | ∞ |

**Table 2.3: PCR conditions for Gibson cloning of MurM.** Conditions used for amplification of gBlocks® gene fragments and the pET22b vector.

Primers were designed using NEBuilder Assembly Tool and gBlocks® were cloned into pET22(b) using the NEBuilder HiFi DNA Assembly Master Mix in accordance with the manufacturers protocol.

Following assembly, the reaction mix was used to directly transform NEB-5$\alpha$ competent cells (Section 2.4.2). A colony PCR (Section 2.3.2) was conducted

and products analysed were using a 1 % (w/v) agarose gel (Section 2.3.3). Plasmid stocks, from positive colonies, colonies were prepared (Section 2.3.7) and sequenced. Constructs created using the Gibson method will be referred to as pET22b::gMurM$_{Pn16}$ and pET22b::gMurM$_{159}$ and the proteins expressed from these will be referred to as gMurM$_{Pn16}$ and gMurM$_{159}$ respectively from hereon in.

### 2.3.9 Site-directed mutagenesis of gMurM$_{Pn16}$

Primers were designed using the NEBaseChanger (Version 1.2.7) and synthesised by IDT. Site directed mutagenesis was conducted using Q5® Site-Directed Mutagenesis Kit as per the manufacturers protocol. In addition 3 % (v/v) DMSO was added to the PCR reaction and the cycling conditions were modified slightly (Table 2.4). Reactions were directly transformed into NEB®5-$\alpha$ competent cells (Section 2.4.2), from which plasmid stocks were miniprepped, quantified and sequenced (Section 2.3.7, 2.3.6 and 2.3.1).

| **Component** | **Temperature (°C)** | **Time (min:sec)** |
|---|---|---|
| Initial Denaturation | 98 | 00:30 |
| Start Cycle (x 16) | | |
| Denaturation | 98 | 00:10 |
| Annealing | 61 | 00:20 |
| Extension | 72 | 05:00 |
| End Cycle | | |
| Final Extension | 72 | 02:00 |
| Hold | 4 | ∞ |

**Table 2.4: PCR conditions for site directed mutagenesis of gMurM$_{Pn16}$**

## 2.4 Protein expression

During cell culture, the optical density at 600 nm ($OD_{600nm}$) was monitored using a Jenway 6306 UV-visible spectrophotometer.

All competent cell lines used throughout this project are detailed in Table 2.5. MurG, porcine heart isocitrate dehydrogenase (IDH), yeast inorganic pyrophosphatase (IPP) and DacB proteins used throughout this project were provided by Julie Tod and Anita Catherwood (University of Warwick).

| ***E. coli* strain** | **Genotype** | **Use** |
|---|---|---|
| NEB®5-$\alpha$ (New England BioLabs, Germany) | *fhuA2* Δ*(argF-lacZ)U169 phoA glnV44* Φ*80* Δ*(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17* | plasmid production |
| BL21(DE3) (Studier and Moffatt, 1986) | F$^-$*ompT hsdSB*(rB$^-$, mB$^-$) *gal dcm* (DE3) | protein expression |
| BL21(DE3)*R (Studier and Moffatt, 1986) | F$^-$*ompT hsdSB*(rB$^-$, mB$^-$) *gal dcm* (DE3) pLysSRARE[*T7p20 ileX argU thrU tyrU glyT thrT argW metT leuW proL* $ori_{p15A}$](Cm$^R$) | protein expression |
| JM109 | *endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB*$^+$ Δ*(lac-proAB) e14- [F' traD36 proAB*$^+$ *lacI*$^q$ *lacZ*Δ*M15] hsdR17*($r_K^- m_K^+$) | protein expression |

**Table 2.5: Competent bacterial strains used for cloning and protein expression.**

### 2.4.1 Competent cell preparation

Competent cell glycerol stocks were streaked for single colonies on LBA plates and incubated at 37 °C overnight. A single colony was inoculated into 2.5 mL of LB media and incubated at 37 °C, shaking at 180 rpm overnight (overnight culture). 250 mL of LB media (supplemented with 20 mM magnesium sulphate ($MgSO_4$)) was inoculated with 2.5 mL of overnight culture and incubated at 37 °C, shaking at 180 rpm. At an $OD_{600nm}$ of 0.4 - 0.6 absorbance units (AU) was achieved. Cells were pelleted by centrifugation at 4,500 *xg* for 5 mins at 4 °C. Cells were kept on ice and all equipment was pre-chilled for remaining manipulations. The cell pellet was gently re-suspended in 100 mL ice-cold 30 mM potassium acetate, 10 mM calcium chloride ($CaCl_2$), 50 mM manganese chloride ($MnCl_2$), 100 mM rubidium chloride ($RbCl_2$) (pH 8.0), 15 % (v/v) glycerol and incubated for 5 mins. Cells were pelleted by centrifugation at 4,000 *xg* for 5 mins at 4 °C. The cell pellet was gently re-suspended in 10 mL of 10 mM 3-(N-morpholino)propanesulfonic acid (MOPS), 74 mM $CaCl_2$, 10 mM $RbCl_2$ (pH 6.5), 15 % (v/v) glycerol and incubated for 15-60 mins. 50 $\mu$L aliquots were flash-frozen in liquid nitrogen and stored at -80 °C.

### 2.4.2 Bacterial transformation

50 $\mu$L aliquots of competent cells were transformed with 50 - 100 ng plasmid, mixed and left on ice for 30 mins. Competent cells were heat shocked at 42 °C for 45 seconds (s) and returned to ice for 2 mins. 200 $\mu$L of SOC was added and cells left to recover at 37 °C, 180 rpm for 60 mins. 50 $\mu$L of recovered cells were plated onto LB + appropriate antibiotic and incubated at 37 °C overnight.

### 2.4.3 Starter culture preparation

For use in trial expressions, a single colony was used to inoculate 3 mL LB media + appropriate antibiotic and incubated at 37 °C, shaking at 180 rpm for 5.5 hours.

For use in large scale expressions, a single colony was used to inoculate 200 mL of LB media + appropriate antibiotic and incubated at 37 °C, shaking at 180 rpm, overnight.

### 2.4.4 Expression trials

Variables that may be altered during an expression trial include; cell line; growth media; expression temperature and time; and concentration of isopropyl $\beta$-D-1-thiogalactopyranoside (IPTG).

100 $\mu$L of starter culture was used to inoculate 10 mL of LB or AIM + appropriate antibiotic. Cultures were incubated at 37 °C, shaking at 180 rpm. When the $OD_{600nm}$ reached 0.6 AU, the temperature was reduced to the desired temperature (15 °C upwards), in addition, cells in LB media were induced by the addition of 0.5 mM IPTG. Cultures were incubated at the desired temperature, for the desired time, shaking at 180 rpm. 1 mL samples were taken at regular time points for analysis.

Time course samples were harvested by centrifugation at 13,000 rpm for 5 mins. Pellets were resuspended in 50 mM sodium phosphate, 300 mM NaCl (adjusted to pH 7.5 using sodium hydroxide (NaOH)) and sonicated twice for 5 s at 10 % power using a Bandelin Sonoplus sonicator. Cell debris was pelleted by centrifugation at 16,000 (*xg*) for 5 mins. The supernatant was retained (soluble protein fraction) and the pellet was resuspended in 500 $\mu$L of 0.1 % (v/v) Tween 20, 1 mM EDTA (pH 8.0) and centrifuged at 16,000 *xg*. The supernatant was discarded and pellet resuspended in 150 $\mu$L of 0.1 % (v/v) Tween 20, 1 mM EDTA (pH 8.0) (insoluble protein fraction).

The soluble and insoluble fractions were analysed by SDS-PAGE (Section 2.5.5).

### 2.4.5 Large scale protein expression

10 mL of a 200 mL starter culture was used to inoculate 1 L LB media supplemented with 0.2 % (v/v) glucose and appropriate antibiotics in 2 L baffled flasks. The cultures were incubated at 37 °C, shaking at 180 rpm. At an $OD_{600nm}$ of 0.6 AU, protein expression was induced by the addition of 0.5 mM IPTG with a concurrent reduction in temperature to 25 °C, cultures were incubated under these conditions for the desired expression period (Section 2.4.6). Cells were pelleted at 10,000 *xg* for 20 mins and stored at -80 °C overnight.

### 2.4.6 Specific protein expression conditions

Following IPTG induction, AlaRS, AlaRS(mutant) and SerRS were expressed for 3 hours whilst Mur ligases (A-F), ThrRS and $gMurM_{159}$ were incubated for 4 hours.

## 2.5 Protein purification

### 2.5.1 Standard protein purification methods

Unless otherwise stated purifications were conducted using an AKTA™ pure protein purification system (GE Healthcare) with UNICORN software (Version 6.3), absorbance at 280 nm was monitored for identification of protein peaks. Cleaning, stripping and storage of chromatography columns was conducted in accordance with the manufacturers recommendations.

#### Immobilised metal affinity chromatography

Immobilised metal affinity chromatography (IMAC) was conducted at room temperature (RT) and used to purify proteins containing C- or N-terminal 6-His tags. A 5 mL HisTrap™ HP column (GE Healthcare) or a 30 mL TALON® Superflow™ (Clontech) column (poured and packed in house as per the

manufacturers recommendations), were used throughout these methods. Columns were equilibrated in 10 column volumes (CV) 0.2 $\mu$m filter sterilised MilliQ $H_2O$, followed by 10 CV of Buffer A (containing no/low Imidazole). Protein samples were loaded directly onto the column via a sample loading loop (unless protein binding was poor, in which case, the sample was allowed to recirculate over the column using a P-1 peristaltic pump (Pharmacia Biotech)) at 4 °C. The column was washed with 10 CV Buffer A and eluted using either a continuous gradient or step-wise increases of Buffer B (high imidazole). Equilibration and elution were conducted at a flow rate of $\leq$ 4 mL.min$^{-1}$, whilst sample loading was conducted at a flow rate of 1 mL.min$^{-1}$. The maximum back pressure for all steps was 0.3 MPa and the sample loaded did not exceed the binding capacity of 200 mg His-tagged protein.

**Size exclusion chromatography**

Size exclusion chromatography (SEC) was used at RT to separate proteins based on size. A Superdex S-200 column (26/60) (poured and packed in house as per the manufacturers recommendations) or a HiPrep$^{TM}$26/60 Sephacryl$^{TM}$ S-200 HR column (GE Healthcare) were used throughout these methods. Columns were equilibrated in 1 CV $H_2O$ and 1 CV equilibration buffer. Sample was loaded in $\leq$ 2 mL using a sample loading loop. A flow rate and maximum pressure of 1 mL.min$^{-1}$ and 0.5 MPa and $\leq$ 1.3 mL.min$^{-1}$ and 0.15 MPa were used for Superdex 200 and Sephacryl S-200 respectively.

### 2.5.2 Specific protein purification methods

**Large scale expression and purification of Mur ligases**

Pellets were thawed on ice and resuspended in 3 mL.g$^{-1}$ 50 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), 1 mM $MgCl_2$ (pH 7.5), 2 mM $\beta$-mercaptoethanol, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin and 0.2 mM

phenylmethane sulfonyl fluoride (PMSF) and 2.5 mg.mL$^{-1}$ lysozyme on a rolling platform for 10 mins at 4 °C. The sample was sonicated on ice for 10 times for 15 s at 70 % power using a Bandelin Sonoplus sonicator, and centrifuged at 10,000 *xg* for 15 mins. The supernatant was subsequently centrifuged at 50,000 *xg* for 45 mins. A 5 mL HiTrap IMAC HP column (GE Healthcare) was equilibrated in 50 mM HEPES, 1 mM $MgCl_2$, 500 mM NaCl (pH 7.5), 5 % (v/v) glycerol, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin and 0.2 mM PMSF. A gradient elution was performed between 0 and 500 mM imidazole in a buffer composed of 50 mM HEPES, 1 mM $MgCl_2$, 500 mM NaCl (pH 7.5), 5 % (v/v) glycerol, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin and 0.2 mM PMSF over 60 mins. Fractions were analysed by SDS-PAGE and fractions containing the protein of interest were pooled and dialysed against 50 mM HEPES, 1 mM $MgCl_2$, 50 mM NaCl (pH 7.5), 3 mM 1,4-dithiothreitol (DTT), 50 % (v/v) glycerol, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin and 0.2 mM PMSF overnight at 4 °C (Section 2.5.3).

**Large scale expression and purification of aminoacyl-tRNA synthetases**

Pellets were thawed on ice and resuspended in 3 mL.g$^{-1}$ 50 mM HEPES, 500 mM NaCl (pH 7.0), 1 mM $\beta$-mercaptoethanol, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin, 0.2 mM PMSF and 2.5 mg.mL$^{-1}$ lysozyme on a rolling platform for 30 mins at 4 °C. Sample was sonicated 10 times for 15 s at 70 % power using a Bandelin Sonoplus sonicator, and centrifuged at 10,000 *xg* for 20 mins. The supernatant was subsequently centrifuged at 50,000 *xg* for 45 mins.

A 5 mL HiTrap IMAC HP column was equilibrated in 50 mM HEPES, 500 mM NaCl, 10 mM Imidazole (pH 7.0), 1 mM $\beta$-mercaptoethanol, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin, 0.2 mM PMSF. Supernatant was loaded using a peristaltic pump and purified using stepwise elutions of 50 mM HEPES, 500 mM NaCl (pH 7.0), 1 mM $\beta$-mercaptoethanol, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin, 0.2 mM PMSF containing 50, 100, 200 and 500 mM imidazole. Fractions were analysed by SDS-PAGE, appropriate fractions were pooled and concentrated using a Vivaspin®20 10 kDa molecular weight cut off (MWCO) (Sartorius).

Superdex 200 (26/60) SEC was implemented for additional purification of AlaRS(mut), SerRS and ThrRS. The column was equilibrated in 50 mM HEPES, 500 mM NaCl (pH 7.0), 1 mM $\beta$-mercaptoethanol, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin, 0.2 mM PMSF, and sample was loaded via a sample loading loop. Fractions were analysed by SDS-PAGE, and fractions containing the protein of interest were pooled and dialysed against 50 mM HEPES, 100 mM NaCl, 1 mM $MgCl_2$ (pH 7.0), 1 mM DTT, 1 $\mu$M pepstatin, 1 $\mu$M leupeptin and 50 % (v/v) glycerol.

**Protein purification of gMurM$_{159}$**

During purification of this protein, all purification steps using an AKTA$^{TM}$ pure protein purification system (GE Healthcare) were conducted in a cold cabinet at 4 °C.

Cells were resuspended in 3 mL.g$^{-1}$ HEPES, 1 mM $MgCl_2$ (adjusted to pH 7.0 with NaOH), 0.2 mM PMSF, 1 $\mu$M leupeptin and 1 $\mu$M pepstatin, and 2.5 mg.mL$^{-1}$ lysozyme at 4 °C for 30 mins. Sample was sonicated 10 x 15 s at 70 % power and centrifuged at 10,000 *xg* for 30 mins.

The pellet was resuspended in 50 mM sodium phosphate, 1 M NaCl, 0.5 mM ethylene glycol-bis($\beta$-aminoethyl ether)-N,N,N,N-tetraacetic acid (EGTA) (adjusted to pH 7.0 with NaOH), 2 mM $\beta$-mercaptoethanol, 0.2 mM PMSF, 1 $\mu$M leupeptin, 1 $\mu$M pepstatin and incubated on a rolling platform for 30 mins at 4 °C to solubilise MurM. The sample was centrifuged at 100,000 *xg* for 45 mins, at 4 °C and supernatant retained. The solubilisation proceedure was repeated with the pellet, and supernatants from each extraction were pooled. Supernatant was sequentially fractionated between 25 and 50 % saturation ammonium sulfate and centrifuged at 50,000 *xg* for 20 mins at 4 °C. The pellet was resuspended in 50 mM sodium phosphate, 500 mM NaCl (adjusted to pH 7.0 with NaOH), 0.2 mM PMSF, 1 $\mu$M leupeptin and 1 $\mu$M pepstatin and purified by SEC on a HiPrep$^{TM}$ 26/60 Sephacryl$^{TM}$ S-200 HR column (GE Healthcare). Fractions were

analysed by SDS-PAGE and fractions containing MurM were pooled. The sample was subsequently loaded, via a sample loading loop, onto a 30 mL TALON® Superflow™ column, equilibrated in IMAC 50 mM sodium phosphate, 500 mM NaCl (pH 7.0), 0.2 mM PMSF, 1 $\mu$M leupeptin and 1 $\mu$M pepstatin. A gradient elution was performed between 0 and 500 mM imidazole in a buffer composed of 50 mM sodium phosphate, 500 mM NaCl, 500 mM Imidazole (adjusted to pH 7.0 with NaOH), 0.2 mM PMSF, 1 $\mu$M leupeptin and 1 $\mu$M pepstatin over 10 CV. Fractions were analysed by SDS-PAGE, and fractions containing MurM were pooled and dialysed against 50 mM HEPES, 1 mM $MgCl_2$, 500 mM NaCl (pH 7.5), 2 mM $\beta$-mercaptoethanol, 0.2 mM PMSF, 1 $\mu$M leupeptin and 1 $\mu$M pepstatin overnight at 4 °C and subsequently against 50 mM HEPES, 1 mM $MgCl_2$, 500 mM NaCl (pH 7.5), 3 mM DTT, 0.2 mM PMSF, 1 $\mu$M leupeptin, 1 $\mu$M pepstatin, 50 % (v/v) glycerol overnight at 4 °C (Section 2.5.3).

### 2.5.3 Protein dialysis and storage

Prior to use, 15.9 mm diameter 12-14 kDa MWCO (Medicell International) dialysis tubing was boiled for 30 mins in 2 % (w/v) sodium bicarbonate and 10 mM EDTA (pH 7.5). Tubing was stored in 10 mM EDTA, 0.5 g sodium azide at 4 °C. Tubing was rinsed with MilliQ $H_2O$ immediately before use.

Following purification, protein samples were dialysed overnight at 4 °C against storage buffer and where necessary concentrated in a Vivaspin®20 10 kDa molecular weight cut off (MWCO) (Sartorius). The concentration of protein samples was determined (Section 2.5.4) and samples were aliquotted and stored at -80 °C.

### 2.5.4 Protein quantification

In a plastic, semi-micro cuvette, 2 $\mu$L of protein was added to diluted Bio-Rad reagent (as per manufacturers instructions), mixed and left to stand for 5 mins. Absorbance at 595 nm was measured using a Jenway 6306 UV-visable spectrophotometer.

Sample was diluted/concentrated as appropriate, until readings were within the linear range. All concentrations were determined in triplicate.

A standard curve was generated using bovine serum albumin (BSA) and used to generate the following formula for calculating the protein concentration:

$$[\text{Protein}]\ (\text{mg.mL}^{-1}) = A_{595\text{nm}} \times 9.75 \times \text{dilution factor}$$

Protein extinction coefficient, molecular weight and pI were calculated using Expasy ProtParam tool (Gasteiger *et al.*, 2005).

### 2.5.5 Protein visualisation

To determine the protein purity following each purification step and at the end of the purification process, protein samples were visualised by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Discontinuous SDS-PAGE (Williams and Reisfeld, 1964; Laemmli, 1970) was used to separate proteins based on molecular weight (MW) under denaturing conditions. Unless otherwise stated, a 12 % resolving gel and a 4 % stacking gel were used for analysis of proteins. Gels were cast using a Mini-PROTEAN® Tetra handcast system, the resolving gel was poured first (3/4 full) and overlaid with ethanol and the gel was allowed to set. The ethanol was removed and a stacking gel was poured, a comb inserted and was allowed to set. Resolving gel comprised 375 mM Tris (pH 8.8), 0.4 % (w/v) SDS, 12 % (w/v) acrylamide:bis acrylamide (37.5:1) polymerised with 0.2 % (w/v) ammonium persulphate (APS) and 0.2 % N,N,N′,N′-Tetramethylethylenediamine (TEMED). Stacking gel comprised 125 mM Tris (pH 6.8), 0.4 % (w/v) SDS, 4 % acrylamide:bis-acrylamide (37.5:1) polymerised with 0.25 % and 0.5 % TEMED.

For analytical gels, loading was standardised to 15 $\mu$g for crude extracts and 10 $\mu$g for pure protein per well (maximum volume per well where protein concentration

was limiting). Samples were prepared in loading buffer comprising 50 mM Tris (pH 6.8), 0.1 % (w/v) SDS, 10 % (v/v) glycerol and bromophenol blue and heat denatured at 98 °C for 10 mins, immediately before loading.

Gels were assembled in a Mini-PROTEAN® Tetra Cell, samples were loaded to each well and gels run at 170 V for 40 mins in running buffer comprising 25 mM Tris, 19 mM Glycine and 17.5 mM SDS (pH 8.3).

Gels were stained for 1 hour using Instant Blue Stain (Expedeon), washed briefly in MilliQ $H_2O$ and imaged using Syngene G:Box with GeneSnap software (Version 7.12).

**Silver staining**

Gel was fixed using 50% (v/v) acetone, 1.2 % (v/v) trichloroacetic acid (TCA) and 0.015 % (v/v) formaldehyde for 15 mins. Gel was washed 3 x in $dH_2O$, incubated in $dH_2O$ for a further 5 mins and washed 3 x in $dH_2O$ again. The gel was soaked in 50 % (v/v) acetone for 50 mins and subsequently in 1 mM sodium thiosulphate ($Na_2S_2O_3$) for 5 mins. Gel was washed 3 x in $dH_2O$ and stained using 15 mM silver nitrate ($AgNO_3$), 0.37 % formaldehyde for 8 mins. The gel was washed 2 x in $dH_2O$ and developed for 10-20 s (or until visible) in 0.7 mM sodium carbonate ($Na_2CO_3$), 1 mM $Na_2S_2O_3$ and 0.015 % formaldehyde. The reaction was quenched by addition of 1 % acetic acid, rinsed in $dH_2O$ and imaged.

**Western blots**

Where required, the presence of $His_6$-tagged proteins were confirmed by Western blot. BioRad Mini-PROTEAN®TGX Stain-Free™ pre-case gels were used for SDS-PAGE.

Proteins were transferred from the gel onto a PVDF membrane using the Trans-Blot®Turbo™ transfer system. The membrane was incubated in 10 % milk

in PBS overnight at 4 °C on a rotating platform. The primary antibody (mouse anti-$His_6$ (Roche)) at 0.2 $\mu$g.mL$^{-1}$ in 0.1 % milk in PBS was added to the membrane and incubated overnight at room temperature. The membrane was washed in 0.1 % tween in PBS, 3 x for 20 mins. The secondary antibody (goat anti-mouse IgG horse radish peroxidase conjugate (Promega)) was diluted 1:2500 in 0.1 % milk in PBS, added to the membrane and incubated for 1-2 hours. The membrane was washed in 0.1 % tween in PBS, 3 x for 20 min. The western blot was developed using BioRad Clarity Western ECL substrates were used at 1:1 ratio and imaged on the Image Quant LAS4000.

## 2.6 Pentapeptide intermediate synthesis and purification

UDP-Mur*N*Ac-pentapeptide (UDP-Mur*N*Ac-5P) intermediates were synthesised as described by Lloyd *et al.* (2008). D-Glu was replaced with D-isoGln for the synthesis of the iGln UDP-Mur*N*Ac-5P variant.

### 2.6.1 Pentapeptide intermediate synthesis reactions

2 mL reactions comprising 50 mM HEPES, 10 mM $MgCl_2$ in 200 mM phosphoenolpyruvate (PEP) (pH 7.6), 1.2 $\mu$M MurA, 57.3 $\mu$M MurB, 1.47 U.mL$^{-1}$ IDH, 5.53 U.mL$^{-1}$ pyruvate kinase (PK), 4.3 $\mu$M MurC, 1.1 $\mu$M MurD, 4 $\mu$M MurE, 8.5 $\mu$M MurF, 1 mM DTT, 50 mM KCl, 0.2 mM nicotinamide adenine dinucleotide phosphate (NADP), 8.2 mM UDP-GlcNAc, 26 mM D-L isocitrate, 6 mM adenosine 5'-triphosphate (ATP), 35 mM L-Ala, 55 mM D-Glu/105 mM D-iso-Gln (iGln), 35 mM L-Lys and 35 mM D-Ala-D-Ala were incubated at 37 °C overnight.

### 2.6.2 Pentapeptide intermediate purification methods

Anion exchange chromatography was used to separate pentapeptide intermediates based on charge. Unless stated otherwise purifications were conducted using

an AKTA$^{TM}$ pure protein purification system (GE Healthcare) with UNICORN software (Version 6.3). Cleaning, stripping and storage of chromatography columns was conducted in accordance with the manufacturers recommendations.

A 75 mL Source 30Q column (poured and packed in house as per the manufacturers recommendations) was used for purification of pentapeptide intermediates. The column was equilibrated in 1.5 CV $H_2O$, 1.5 CV 1 M ammonium acetate (adjusted to pH 7.6 with acetic acid ($CH_3COOH$) and 1.5 CV 10 mM ammonium acetate (adjusted to pH 7.6 with $CH_3COOH$). A flow rate and maximum pressure of 10 mL.min$^{-1}$ and 0.5 MPa respectively were used throughout.

Sample was centrifuged at 1800 *xg* for 45 mins in a Vivaspin®20 centrifugal concentrator. The sample was loaded on to a Source 30Q anion exchange column manually via the system pump, washed with 100 mL of 10 mM ammonium acetate and eluted with a linear gradient from 0.01-1 M ammonium acetate (adjusted to pH 7.6 with $CH_3COOH$) over 120 mins. Absorbance at 280 nm, 254 nm and 218 nm was monitored for identification of the uridine containing pentapeptide peak. Fractions containing the pentapeptide were pooled, lyophilised and resuspended in 500 mL $H_2O$ at least six times using an ALPHA 2-4 LD PLUS Freeze Drier (CHRIST). The product was finally resuspended in 1 mL $H_2O$ and stored at -80 °C

### 2.6.3 Pentapeptide intermediate quantification

180 $\mu$L sterile $H_2O$ added to a quartz cuvette and blanked on the Jenway 7305 UV-visible spectrophotometer at an absorbance of 260 nm. 2 $\mu$Lof purified UDP-Mur*N*Ac-5P sample was added and mixed thoroughly by pipetting (if sample was highly concentrated it was diluted as required). This was conducted in triplicate, an average reading for absorbance at 260 nm was used to calculate the concentration using the following formula:

$$\text{Concentration (M)} = \frac{\text{A}_{260\text{nm}}}{\epsilon \text{ of uridine}} \times \frac{\text{total cuvette volume}}{\text{sample volume} \times \text{dilution factor}}$$

where the extinction coefficient ($\epsilon$) at 260 nm of uridine was 10,000 $M^{-1}.cm^{-1}$.

### 2.6.4 Pentapeptide intermediate activity determination

A 200 $\mu$L reaction comprising 50 mM HEPES, 10 mM $MgCl_2$ (pH 7.0), 270 $\mu$g DacB, 2.5 $\mu$M amplex red, 10 $\mu$g horse radish peroxidase and 66 $\mu$g D-amino acid oxidase was initiated by addition of pentapeptide intermediate and monitored at 555 nm using a Varian Cary 100 spectrophotometer at 37 °C. The rate of reaction varies depending upon the volume and concentrations of pentapeptide added to the reaction. A presence of any rate is considered a positive result, and the peptide intermediate can be used for downstream synthesis reactions.

### 2.6.5 Intermediate purity determination

A MonoQ (5/50) column (GE Healthcare) was used for analysis of the purity of pentapeptide intermediates. The column was equilibrated in 3 CV $H_2O$, 3 CV 1 M ammonium acetate (adjusted to pH 7.6 with $CH_3COOH$) and 3 CV 10 mM ammonium acetate (adjusted to pH 7.6 with $CH_3COOH$). A flow rate and maximum pressure limit of 0.5 $mL.min^{-1}$ and 3 MPa were used throughout. 1 mL of pentapeptide sample was diluted to $\leq$ 300 $\mu$M in 10 mM ammonium acetate (adjusted to pH 7.6 with $CH_3COOH$). Sample was loaded to the MonoQ (5/50) column via a sample loading loop and eluted using a linear gradient of 0-100 % 1 M ammonium acetate (adjusted to pH 7.6 with $CH_3COOH$) over 20 mins. The absorbance at 280 nm and 254 nm was monitored for identification of pentapeptide peaks and the peak integration function in UNICORN software (Version 6.3) was used to determine the purity using the following equation:

$$\text{Purity (\%)} = \frac{\text{area under the UDP-Mur}N\text{Ac-5P peak}}{\text{area under all peaks}}$$

**Intermediate identity confirmation**

The identity of the product was confirmed by negative ion nanospray time-of-flight (TOF) mass spectrometric analysis using a Waters SYNAPT G2-Si mass spectrometer and MassLynx$^{TM}$ software (Version 4.1). The instrument was calibrated with NaI clusters and the sample, in 50 % (v/v) acetonitrile, was introduced via a nanospray needle at a capillary voltage of 1.5 kV. The expected exact mass of UDP-MurNAc-5P and iGln UDP-MurNAc-5P were 1149.35 and 1148.37 respectively.

## 2.7 Lipid substrate synthesis

Throughout this section, where methods are identical, Lipid is used to refer to Lipid II(Lys) or iGln Lipid II(Lys).

Lipid intermediates were synthesised using methods described by Breukink *et al.* (2003). Minor modifications (detailed below) were used to maximise yield and allow purification of the iGln Lipid II(Lys) variant.

Solvent A (SolA) comprising chloroform:methanol:water (2:3:1) was used throughout this section.

### 2.7.1 Preparation of *Micrococcus flavus* membranes

*Micrococcus flavus* glycerol stocks were streaked onto TSB agar plates and incubated at 37 °C overnight. A single colony was inoculated into 100 mL of TSB media and incubated at 37 °C, shaking at 180 rpm overnight. 8 x 650 mL of TSB supplemented with 0.1 % (v/v) glucose, in 2 L baffled flasks, was inoculated with 6.5 mL starter culture and incubated at 37 °C, shaking at 180 rpm until an $OD_{600nm}$ of 9.0 AU was achieved. Cells were pelleted at this point (or earlier if they are no longer in exponential growth phase) by centrifugation at 8,000 *xg* for 20 mins. The pellet was resuspended in 25 mL of 20 mM Tris, 1 mM $MgCl_2$ (pH 7.5), 2 mM

$\beta$-mercaptoethanol and centrifuged at 15,000 *xg* for 20 mins. Supernatant was discarded and pellet was stored at -20 °C.

Pellet was resuspended in 3 mL.g$^{-1}$ of 20 mM Tris, 1 mM $MgCl_2$ (pH 7.5), 2 mM $\beta$-mercaptoethanol plus 2.5 mg.mL$^{-1}$ lysozyme and incubated on a rolling platform at 4 °C for 20 mins. Sample was passed through a continuous cell disrupter (Constant Cell Disruption Systems) at 30 kspi twice and centrifuged at 10,000 *xg* for 1 hour. Pellets were resuspended in 3 mL of supernatant, aliquoted and stored at -80 °C.

### 2.7.2 Optimisation of iGln Lipid II(Lys) synthesis

The work outlined in this section was conducted in collaboration with Catherine Rowland (PhD Student).

#### Small scale trial syntheses

Reactions comprising 0.23 mg undecaprenyl phosphate, 2 mM or 4 mM iGln UDP-MurNAc-5P(lys), 6 mM UDP GlcNAc, 0.57 mg.mL$^{-1}$ MurG, 5 mM $MgCl_2$ and 0 $\mu$M or 5 $\mu$M moenomycin were made up to 200 $\mu$L with *M. flavus* membranes, and incubated at 37 °C for 3 hours or overnight. Using this reaction as a template, variation of the concentration of iGln UDP-5P(lys), the presence of moenomycin and incubation length were tested for their impact on synthesis of the desired iGln Lipid II(Lys). Samples were isolated and analysed by TLC as described in section 2.7.4 and 2.7.5 respectively.

#### Large scale time course experiments

2 mL reactions using the most promising conditions as established in Section 2.7.2 were tested on a large scale. 200 $\mu$L samples were taken and frozen at -20 °C every

hour for 12 hours. Samples were isolated and analysed by TLC as described in Sections 2.7.4 and 2.7.5 respectively.

### 2.7.3 Lipid synthesis reactions

In a glass vial, 2 mg undecaprenyl phosphate in 100 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 1 % (w/v) Triton X-100 with 2 mM UDP-MurNac-5P or 4 mM iGln UDP-MurNAc-5P, 6 mM UDP-GlcNAc, 0.57 $mg.mL^{-1}$ MurG and 5 mM $MgCl_2$ were made up to 2 mL with *M. flavus* membranes and incubated at 37 °C overnight.

### 2.7.4 Isolation and purification of Lipid

#### Isolation of Lipid

Lipid synthesis reactions were terminated by the addition of 1 x incubation volume of 6 M pyridinium acetate and 2 x incubation volume of N-butanol. Samples were voretexed and centrifuged at 3,000 *xg* for 10 mins to seperate organic and aqueous phases. The organic (top) phase was recovered and re-extracted in an equal volume of $H_2O$. The sample was vortexed and centrifuged at 3,000 *xg* for 10 mins, the organic (top) phase was dried down using a Rotavapor R-210 (or Vacuum Pump V-710 (Buchi) for small sample volumes). Dried samples were resuspended in 1.5 x CV of SolA.

#### Column preparation

For a full scale synthesis (2 mL) a bed volume of 4 mL was used, this was scaled accordingly for smaller syntheses. A diethylaminoethyl (DEAE)-sephacel (acetate counter ion) gravity flow column was prepared using a glass burette containing a small amount of glass wool in the tip, as a filter. The column was equilibrated with 10 CV 1 M ammonium acetate, 15 CV $H_2O$ and 10 CV Sol A.

**Purification of Lipid**

The dried sample was resuspended in 1.5 x CV of Sol A. All fractions were collected in glass bijoux tubes for analysis. Sample was loaded to the column and 3 CV SolA used to wash the column. The Lipid was eluted using stepwise washes of 12 mL chloroform:methanol:aqueous ammonium bicarbonate (2:3:1), whereby the concentration of the ammonium bicarbonate component increased with every wash. Washes whereby the aqueous component of the eluant contained 50, 75, 100, 150, 200, 250, 300, 500 mM and 1 M ammonium bicarbonate were used for Lipid II(Lys) purifications. Additional washes where the aqueous component of the eluant contained of 60, 65, 70, 80, 85, 90, 95 mM ammonium bicarbonate were included for the purification of iGln Lipid II(Lys). 400 $\mu$L samples from each fraction were analysed by TLC (2.7.5) and those containing product were pooled and dried under vacuum. The dried sample was resuspended in $H_2O$ and lyophilised at least three times using a ALPHA 2-4 LD PLUS Freeze Drier (CHRIST). The sample was resuspended in 1 mL SolA, and stored at -80 °C.

### 2.7.5 Visualisation of Lipid fractions

Fractions were analysed by thin layer chromatography (TLC). 400 $\mu$Lsamples were resuspended in 50 $\mu$L Sol A, 25 $\mu$L was spotted onto a silica 60-coated ALUGRAM Xtra SIL G plates (Macherey-Nagel) and the bottom $\sim$2 cm of the plate was submerged in chloroform/methanol/water/ammonia (88:48:10:1) for $\sim$2.5 hours (or until the solvent front was $\sim$2 cm from the top). The TLC plate was air dried, stained by exposure to iodine vapour for 20 mins and visualised on a flatbed scanner.

### 2.7.6 Quantification of Lipid

The Lipid concentration was determined by acid hydrolysis of the phosphate moiety and detection using a phosphate release assay (Webb 1992).

Two 50 $\mu$L samples of Lipid II and 50 $\mu$L SolA (as a negative control) were dried down using nitrogen gas and resuspended in 50 $\mu$L of 50 mM HEPES, 10 mM $MgCl_2$, 30 mM KCl and 1.5 % (w/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS). 50 $\mu$L of 1M hydrochloric acid (HCl) was added to one Lipid II sample and the SolA negative control and the samples were boiled at 100 °C for 30 mins. 1 M sodium hydroxide (NaOH) was added to adjust the boiled samples to pH 7.6.

A 200 $\mu$L phosphate release assay (Figure 2.1) comprising 50 mM HEPES, 10 mM $MgCl_2$ (pH 7.6), 1 unit purine nucleoside phosphorylase (PNP), 1 unit IPP and 200 $\mu$M 7-methyl-6-thioguanosine (MESG) (Berry and Associates) were initiated with 10 $\mu$L Lipid II sample and the absorbance at 360 nm was monitored using a Varian Cary 100 spectrophotometer at 37 °C. Reactions were conducted in triplicate, and both the non-hydrolysed Lipid II control and the SolA control used to calculate the Lipid II concentration using the following equation:

The $\Delta$ extinction coefficient at 360 nm generated by phosphorolysis of MESG was 10,000 $M^{-1}.cm^{-1}$, per molar phosphate (Webb, 1992). Lipid II concentrations were calculated based on the release of 2 mols of phosphate per mol of Lipid.

### 2.7.7 Lipid identity confirmation

All Lipid mass spectrometry (MS) was conducted by Adrian Lloyd, Anita Catherwood or Julie Tod. Spectra analysis was conducted with assistance from Adrian Lloyd.

The identity of the product was confirmed by negative ion nanospray time of flight (TOF) MS using a Waters SYNAPT G2-Si quadrupole-TOF instrument operating in resolution mode equipped with a nanospray source. The instrument was calibrated with sodium iodide (NaI) clusters over a 200-2500 *m/z* range, with an error of less than 1 ppm. Lipid sample was dried, resuspended at a final concentration of 1-5 mM, in 70 % (v/v) methanol and 30 % (v/v) 25 mM ammonium acetate and introduced

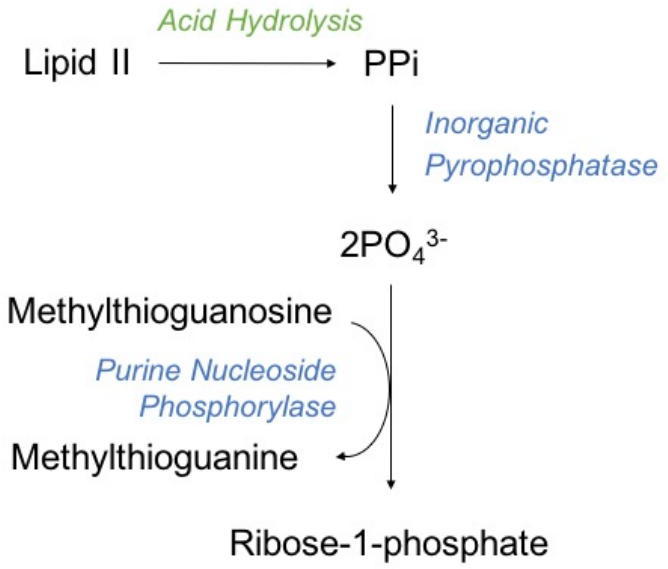



**Figure 2.1: Phosphate release assay.** $PP_i$ produced by the acid hydrolysis of Lipid II is converted to $P_i$ by IPP. PNP catalyses the conversion of $P_i$ and methylthioguanosine to ribose-1-phosphate and methylthioguanine, which absorbs at 360 nm.

into the instrument using a Waters thin wall nanoflow capillary. Data was collected for ∼10 mins at a capillary voltage of 2.0 kV, cone voltage of 100 V and source offset of 41 V. MassLynx$^{TM}$software (Version 4.1) was used to analyse the data.

Collision induced fragmentation of Lipid species of the desired *m/z* was performed in the trap cell of the mass spectrometer. Argon flow was set at 2 ml.min-1 and data was acquired for collision energies of 0, 15, 30, 45, 60, 75 and 90 electronvolts (eV).

## 2.8 tRNA substrate preparation

### 2.8.1 Growth of *S. pneumoniae* (159) and purification of crude tRNA

*S. pneumoniae* (159) glycerol stocks were streaked onto Muller-Hinton agar plates and incubated at 37 °C at 5 % $CO_2$ overnight. Cells were harvested into BHI media to create an inoculum of $OD_{600nm}$ of >1.0 AU. 1 mL of innoculum was added to 1 L BHI + 20-50 $\mu$g catalase and grown anaerobically at 37 °C. The culture at $OD_{600nm}$ was monitored at regular intervals using a Jenway 6306 UV-visable spectrophotometer.

Immediately prior to reaching stationary growth phase, cultures were harvested, by centrifugation at 10,000 *xg* for 20 mins. Pellets were resuspended in 20 mM Tris, 2 mM Magnesium acetate ($Mg(CH_3COO)_2$) (pH 7.4) and incubated at 37 °C, shaking at 130 rpm for 1 hour. tRNA was isolated as detailed by Zubay (1962) and von Ehrenstein (1967).

### 2.8.2 Production of single species of *S. pneumoniae* $tRNA^{Ala}$, $tRNA^{Ser}$ and $tRNA^{Thr}$

The sequences of each $tRNA^{Ala}$, $tRNA^{Ser}$ and $tRNA^{Thr}$ gene from *S. pneumoniae* R6 were identified (Chan and Lowe, 2009). Sequences of different isoacceptors were synthesised and cloned into pIDTSMART-Kan by IDT. A T7 promotor sequence at the 5′ end and a BstNI cleavage site at the 3′ end (to provide a CCA terminus for run off transcription) were incorporated into the gene. Forward and Reverse primers were also synthesised by IDT. The tRNA gene sequences and primers are detailed in Table 2.6.

PCR mixes comprising 1 x Q5 Master Mix, 1.25 $\mu$M Forward Primer, 1.25 $\mu$M Reverse Primer, 40 ng plasmid were used to amplify the tRNA genes using the SureCycler 8800 thermocycler (Agilent Technologies) Touchdown PCR programme. Details of PCR conditions are detailed in Table 2.7. PCR products were analysed on a 2 % (w/v) agarose gel, purified using Monarch® PCR and DNA clean up kit and quantified (Section 2.8.3).

To reveal the 3′ CCA motif, 1 U of BstNI per $\mu$g of tRNA was incubated at 60 °C for 1 hour, as per the manufacturer's protocol. Samples were purified using Monarch® PCR and DNA clean up kit and quantified.

| Gene Name and Locus Tag | Gene Sequence | Iso | Forward Primer | Reverse Primer |
|---|---|---|---|---|
| $tRNA^{Ala4}$, sprt54 | GATCCCAGTAATACGACTCACTATAGGGGCCTTAGCTCAGCTGGGAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCCGCTAGGCTCCACCAGGAATTC | UGC | GATCCCAGTAATACGACTCACTATAGG (1) | TGGTGGAGCCTAGCGGGAT |
| $tRNA^{Ser1}$, sprt06 | GATCCCAGTAATACGACTCACTATAGGAAGATTACTCAAGAGGCTTAAGAGGCCGTGTTGGAAACGCGGTAGGCGTGTAATAGCGTGCGTGGGTTCGAATCCCATGTCTTCCGCCAGGAATTC | GGA | GATCCCAGTAATACGACTCACTATAGG (1) | TGGCGGAAGACATGGGATTC |
| $tRNA^{Ser2}$, sprt12 | GATCCCAGTAATACGACTCACTATAGGAGAATTACTCAAGAGGCTGAAGAGGACGGTTTGCTAAATCGTTAGGTCGGGTAACTGGCGCAAGGGTTCGAATCCCTTATTCTCCGCCAGGAATTC | GCU | GATCCCAGTAATACGACTCACTATAGG (1) | TGGCGGAGAATAAGGGATTCG |
| $tRNA^{Ser3}$, sprt17 | GATCCCAGTAATACGACTCACTATAGGAGGATTACCCAAGTCCGGCTGAAGGGAACGGTCTTGAAAACCGTCAGGCGTGTAAAAGCGTGCGTGGGTTCGAATCCCACATCCTCCTCCAGGAATTC | UGA | GATCCCAGTAATACGACTCACTATAGG (1) | TGGAGGAGGATGTGGGATT |
| $tRNA^{Thr1}$, sprt07 | GATCCCAGTAATACGACTCACTATAGCTGATTTAGCTCAGTTGGCAGAGCGCATCCATGGTAAGGATGAGGTCGCCGGTTCAATCCCGGCAATTAGCACCAGGAATTC | GGT | GATCCCAGTAATACGACTCACTATAGC (2) | TGGTGCTAATTGCCGGGATT |
| $tRNA^{Thr2}$, sprt24 | GATCCCAGTAATACGACTCACTATAGCCGGCTTAGCTCAGTTGGTAGAGCATCTGATTTGTAATCAGAGGGTCGCGTGTTCAAGTCATGTAGCCGGCACCAGGAATTC | TGT | GATCCCAGTAATACGACTCACTATAGC (2) | TGGTGCTAATTGCCGGGATT |

**Table 2.6: tRNA gene sequences to be cloned into pIDTSMART-Kan vectors by IDT**. All tRNA isoacceptors will be referred to as per the gene name. The gene sequence (black), additional 5′ sequence containing a T7 promotor (red) and the additional 3′ sequence containing a BstNI cleavage site (blue) are shown. The anticodon for the specific isoacceptor, and forward and reverse primers are detailed.

| Component | Temperature (°C) | Time (min:sec) |
|---|---|---|
| Initial Denaturation | 98 | 00:30 |
| Max Annealing Temp | • $tRNA^{Ala4}$ - 65<br>• $tRNA^{Ser1}$ - 58<br>• $tRNA^{Ser2/3}$ - 55<br>• $tRNA^{Thr1}$ - 58<br>• $tRNA^{Thr2}$ - 65 | |
| Min Annealing Temp | • $tRNA^{Ala4}$ - 60<br>• $tRNA^{Ser1/2/3}$ - 50<br>• $tRNA^{Thr1}$ - 50<br>• $tRNA^{Thr2}$ - 60 | 00:30 |
| Elongation Temp | 72 | 00:05 |
| Cycle (x35) | | |

**Table 2.7: Touchdown PCR conditions for amplification tRNA genes from pIDTSMART-Kan constructs**

*In vitro* transcription was conducted using TranscriptAid T7 High Yield Transcription Kit (Thermo), as per the manufacturers recommendations. Products were quantified (Section 2.8.3 and analysed for purity Section 2.8.4).

### 2.8.3 Quantification of tRNA

$A_{280nm}$ of a 1 $\mu$L sample was determined using an IMPLEN NanoPhotometer®. The concentration of tRNA was calculated using the following equation:

$$[\text{tRNA}]\ (\mu\text{M}) = \frac{A_{260nm}}{\epsilon} \times 1000$$

Where $\epsilon$ is extinction coefficient, predicted using Oligo Extinction Coefficient Calculator.
(https://www.scripps.edu/researchservices/old/corefac/biopolymercalc2.html).

### 2.8.4 Visualisation of tRNA samples

7 M urea polyacrylamide gel electrophoresis (7M urea PAGE) was used to analyse tRNA purity. Gels were cast using a Mini-PROTEAN® Tetra handcast system (BioRad), the gel was poured, a comb inserted and was allowed to set. The gel comprised 90 mM Tris base, 90 mM boric acid, 10 mM EDTA (adjusted to pH 8.0 with HCl), 7 M urea, 40 % (v/v) AccuGel™29:1 (National Diagnostics), 0.06 % (w/v) APS and 0.006 % (v/v) TEMED.

For analysis of gels, loading was standardised to 20 $\mu$g crude tRNA and 5 $\mu$g *in vitro* transcribed tRNA, 10 $\mu$g Low Molecular Weight DNA Ladder (NEB) was used. Samples were prepared in loading buffer comprising 1 M urea, 90 mM Tris base, 90 mM boric acid, 10 mM EDTA (adjusted to pH 8.0 with HCl), 0.1 % (w/v) bromophenol blue and 0.1 % (w/v) xylene cyanol.

Gels were assembled in a Mini-PROTEAN® Tetra Cell, samples loaded to each well and gels run at 120 V for 3 hours in 90 mM Tris base, 90 mM boric acid, 10 mM EDTA (adjusted to pH 8.0 with HCl).

Gels were stained for 15 mins using 0.4 % (w/v) Toludine Blue in 50 % (v/v) methanol and 10 % (v/v) acetic acid. Gels were de-stained overnight in 50 % (v/v) Methanol and 10 % (v/v) acetic acid, rehydrated in $dH_2O$ and imaged using ImageQuant LAS 4000 instrument (GE Healthcare).

## 2.9 Aminoacyl-tRNA synthetase spectrophotometric assays

All assays were conducted using a Cary 100 Bio UV-Visable Spectrophotometer (Agilent) with Cary WinUV Software (Version 4.20). Kinetic analysis was conducted

using Prism 7 (Version 7.0c).

### 2.9.1 Amino acid activation by aminoacyl-tRNA synthetases

#### Amino acid activation assay

A 200 $\mu$L 5'-adenylyl-$\beta$,$\gamma$-imidodiphosphate (ADPNP) reaction (Figure 2.2) comprising 50 mM HEPES, 10 mM $MgCl_2$ (pH 7.6), 1 mM DTT, 10 mM D-Glucose, 50 mM KCl, 0.15 mM $NADP^+$, 0.34 U Hexokinase (Roche), 0.17 U glucose-6-phosphate dehydrogenase (Roche), 0.5 mM ADPNP, 0.108 mM inorganic pyrophosphate ($PP_i$) and appropriate concentration of aaRS was prepared. The reaction was initiated by addition of varying concentrations of amino acid. Absorbance at 340 nm was measured at 37 °C using a Cary 100 Bio UV-Visable Spectrophotometer (Agilent) with Cary WinUV Software (Version 4.20). The initial (steady state) change in velocity was converted to turnover number related to substrate concentration and plotted using Prism 7 (Version 7.0c).

#### Specific assay conditions

0.8 $\mu$M AlaRS and 0.05-20 mM L-Ala or 20-200 mM L-Ser, was used to measure the rate of amino acid activation by alanyl-tRNA Synthetase (AlaRS). 4.08 $\mu$M ThrRS and 0.5-20 mM L-Thr was used to measure the rate of amino acid activation by threonyl-tRNA synthetase (ThrRS).

### 2.9.2 Pre- and post-transfer editing of aminoacyl-tRNA synthetases

#### Pre and post-transfer editing assay

A pre- and post-transfer editing assay (Figure 2.3) comprised 50 mM HEPES, 10 mM $MgCl_2$ (pH 7.6), 1 mM DTT, 50 mM KCl, 1 mM ATP, 2 mM phophoenolpyruvate (PEP) (MP Biomedicals), 0.3 mM reduced nicotinamide adenine dinucleotide

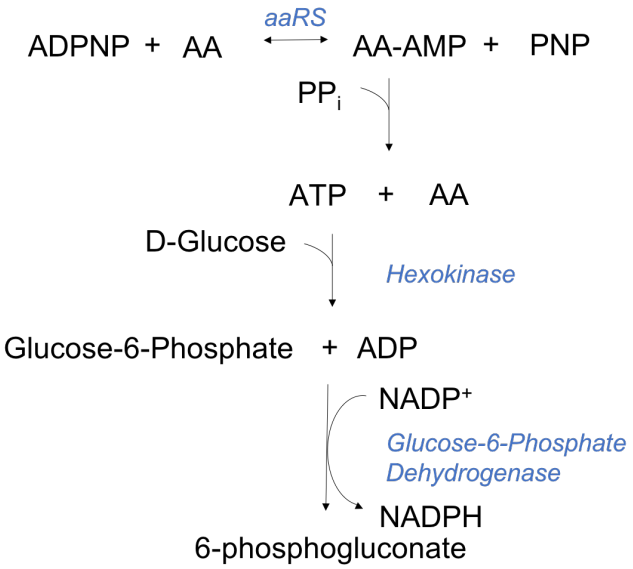



**Figure 2.2: ADPNP assay for measuring enzymatic activity.** ADPNP is an ATP analogue which is consumed during the formation of an aminoacyl-adenylate. The aminoacyl-adenylate is cleaved by $PP_i$ to form ATP. Hexokinase* then utilises the ATP and phosphorylates D-Glucose to form Glucose-6-phosphate. Glucose-6-phosphate dehydrogenase then converts Glucose-6-phosphate to 6-phosphogluconate by the reduction of $NADP^+$ to NADPH. The production of NADPH is measured as an increase in absorbance at 340 nm. *Hexokinase is able to completely discriminate between ATP and ADPNP.

(NADH), 60-120 U myokinase, 1.2-2.0 U pyruvate kinase/1.8-2.8 U lactate dehydrogenase (PK/LDH), and an appropriate concentration of non-congnate amino acid. The pre-transfer reaction was initiated by addition of aaRS, and the post-transfer reaction subsequently initiated by addition of *E. coli* crude tRNA. Absorbance at 340 nm was measured using a Cary 100 UV-Vis Spectrophotometer at 37 °C. The pre-transfer editing rate was subtracted from the rate obtained on addition of tRNA to yield the post-transfer editing rate. Editing rates were plotted against [$H_2O_2$] or time, defending upon the experimental design.

### Effect of $H_2O_2$ concentration on editing of aminoacyl-tRNA synthetases

The effect of $H_2O_2$ on pre- and post-transfer editing rates of AlaRS and ThrRS was established by addition of varying concentrations of $H_2O_2$ to the reaction prior to initiation.

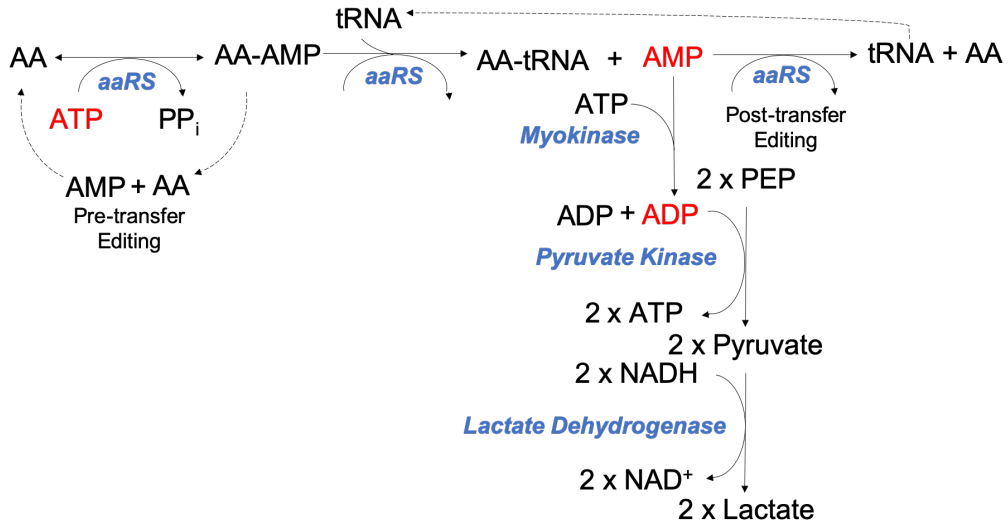



**Figure 2.3: Schematic of spectrophotometric pre- and post-transfer editing assay**. In the presence of amino acid but not tRNA, the ATP is consumed in the formation of aminoacyl adenylate, the rate observed is due to Pre-transfer editing. Upon addition of the tRNA aa-tRNA is formed, this is then edited by the synthetase releasing AMP, which is converted to adenosine 5'diphosphate (ADP) by myokinase and ATP by pyrivate kinase using PEP as a phosphoryl donor. The resulting pyruvate is then converted to lactate by lactate dehydrogenase (LDH) converting NADH into $NAD^+$, producing a fall in the absorbance at 340 nm.

In the case of AlaRS, reactions contained 200 mM L-Serine (or 200 mM L-Ala for control) and 0 - 75 mM $H_2O_2$. Pre-transfer editing was initiated by the addition of 0.3 $\mu$M AlaRS and post-transfer editing was initated by addition of 0.66 $mg.mL^{-1}$ *E. coli* crude tRNA. In the case of ThrRS reactions contained 50 mM L-Ser (or 50 mM L-Thr for control) and 0 - 50 mM $H_2O_2$. Pre-transfer editing was initiated by the addition of 10 $\mu$M ThrRS and post-transfer editing was initiated by addition of 1.9 $mg.mL^{-1}$ *E. coli* crude tRNA.

To provide a negative control, the post-transfer editing rate of the aaRS in the presence of the non-cognate amino acid was established in the absence of $H_2O_2$.

Finally, the effect of $H_2O_2$ on the coupled system was characterized, to ensure that the measurements were a direct result of the effect of $H_2O_2$ on the aaRS.

**Effect of incubation time on pre- and post-transfer editing of aminoacyl-tRNA synthetases**

The effect of time on pre- and post-transfer editing was investigated by incubation of AlaRS of ThrRS with a fixed concentration of $H_2O_2$ for varying lengths of time at room temperature. At specific time points, samples of the aaRS was taken and used to initiate the reaction (as above).

For AlaRS, pre-transfer editing reactions were initiated with 0.3 $\mu$M pre-treated AlaRS after 0-30 mins incubation with 50 mM $H_2O_2$. Post-transfer editing reactions were initiated with 0.76 mg.mL$^{-1}$ *E. coli* crude tRNA. Negative controls were conducted at 0 and 30 min time points.

For ThrRS, pre-transfer editing reactions were initiated with 10 $\mu$M pre-treated ThrRS after 0-40 mins incubation with 20 mM $H_2O_2$. Post-transfer editing reactions were initiated with 0.76 mg.mL$^{-1}$ *E. coil* crude tRNA. Negative controls were conducted at 0 and 40 min time points.

To provide a negative control, aaRS was incubated in the absence of $H_2O_2$, samples were taken at time points and assayed.

Since the reaction master mix contained DTT, a reducing agent, the effect of omitting DTT from the reaction mixture was also determined. aaRS was incubated with $H_2O_2$ for 10 mins, and post-transfer editing was assayed in the presence and absence of DTT.

## 2.10 Investigating aminoacyl-tRNA synthetase oxidation by mass spectrometry

### 2.10.1 Intact protein mass spectrometry

AlaRS and ThrRS were treated with 10,000 x molar ratio $H_2O_2$ or 200 x molar ratio of $H_2O_2$ for 30 mins and 40 mins respectively. AlaRS and ThrRS were purified in HPLC grade $H_2O$ using a PD-10 column (pre-equilibrated as per manufacturers protocol) and analysed by positive ion nanospray TOF MS using a Waters SYNAPT G2-Si quadrupole-TOF instrument with a nanospray source. The instrument was calibrated using 2 $\mu$g.µl$^{-1}$ NaI and the sample prepared in 50 % (v/v) methanol and 0.1 % formic acid (FA) was introduced via a nanospray needle at a capillary voltage of 2 kV. Analysis was conducted using MassLynx$^{TM}$software (Version 4.1 SCN957). Mock spectra were generated to assess the quality of data and the MaxEnt function was used to determine the molecular mass of components present in the sample.

### 2.10.2 Peptide-based mass spectrometry

AlaRS and ThrRS were treated with 10,000 x molar ratio $H_2O_2$ or 200 x molar ratio of $H_2O_2$ for 30 mins and 40 mins respectively. AlaRS and ThrRS were buffer exchanged into 50 mM HEPES, 100 mM NaCl, 1 mM $MgCl_2$ (pH 7.0), 1 $\mu$M pepstatin, 1 $\mu$M leupeptin and 50 % (v/v) glycerol, to remove the $H_2O_2$. SDS-PAGE of samples was conducted and gel bands at the appropriate molecular weight were cut out and diced into 2 mm x 2 mm cubes using a scalpel. The gel was de-stained completely using 50 mM ammonium bicarbonate with 50 % ethanol, and dehydrated using 100 % ethanol for 5 mins. Gel pieces were incubated with 10 mM Tris(2-carboxyethyl)phosphine (TCEP) and 40 mM chloroacetamide (CCA) for 5 mins at 70 °C, and subsequently washed in 50 % ethanol and 50 mM ammonium bicarbonate. Gel was dehydrated for 5 mins in 100 % ethanol and treated with 2.5 ng.µl$^{-1}$ trypsin (diluted in 50 mM ammonium bicarbonate) at 37 °C overnight. Sample was sonicated (in a waterbath)

in 25 % acetonitrile, 5 % formic acid 3 times for 10 mins each. Liquid was retained and pooled, dried down using a speed vac and resuspended to a final volume of 50 $\mu$L in 2 % acetonitrile and 0.1 % trifluoroacetic acid.

Tandem mass spectrometry of samples was conducted by the Proteomics Research Technology Platform (University of Warwick).

Analysis was conducted using Scaffold (Version 4.8.1).

## 2.11 MurM assays

### 2.11.1 Mass spectrometry MurM activity assay

This MurM assay is coupled to aminoacyl-tRNA synthetase activity and Lipid product was detected by mass spectrometry. Reactions contained 50 mM HEPES, 10 mM $MgCl_2$, 30 mM KCl, 1.5 % CHAPS (w/v) (pH 7.6), 5 mM ATP, 2 mM DTT, 500 $\mu$M Lipid II (lys)-5P, 3 $\mu$M AlaRS, 10 mM L-Ala or L-Ser and 1.5 $mg.mL^{-1}$ *S. pneumoniae* (159) tRNA either with or without 1.5 $\mu$M $gMurM_{159}$. Reactions were incubated at 37 °C for 1 hour, purified (using stock ammonium bicarbonate concentrations of 50, 100 mM and 1 M for elution) and freeze dried as per Section 2.7.4. The sample was resuspended in 70 % (v/v) methanol and 30 % (v/v) 25 mM ammonium acetate. The identity of the product was confirmed by negative ion nanospray TOF MS using a Waters SYNAPT G2-Si mass spectrometer. The instrument was calibrated with NaI clusters and the sample was introduced via a nanospray needle at a capillary voltage of 2.0 kV. The expected exact mass of Lipid II(Lys)-Ala/Ser and iGln Lipid II(Lys)-Ala/Ser were 1946.10/1962.09 and 1945.11/1961.11 respectively.

Collision induced dissociation (CID) in combination with positive ion tandem MS (MS/MS) was conducted using the Waters SYNAPT G2-Si mass spectrometer to determine the amino acid sequence and therefore exact location of amino acid addition to the peptide stem of Lipid II(Lys).

## 2.12 Radioactive assays

The experimental set ups were initially tested with readily available crude *E. coli* tRNA and subsequently repeated with the appropriate *S. pneumoniae* tRNA. [$^3$H]L-Ala (Moravek), [$^3$H]L-Ser (Moravek) and [$^3$H]L-Thr (ARC Inc) were all prepared to stock concentrations of 0.6 mM. When experiments were repeated with multiple radioactive amino acids, the amino acid is referred to as [$^3$H]-aa. Samples were spotted onto 1 cm x 1 cm Whatman cellulose chromatography papers 3MM. Radioactivity was followed by scintilltion counting using a Tri-Carb 2800TR Liquid Scintillation Analyzer.

### 2.12.1 Radioactive tRNA-aminoacylation activity assay

Reactions containing 30 mM HEPES, 15 mM $MgCl_2$, 25 mM KCl, 5 mM DTT, 2 mM ATP (pH 7.5), 5 mM DTT, 73 $\mu$g crude OR 0.3 $\mu$M pure tRNA, 7 $\mu$g IPP and 0.1 mM [$^3$H]-aa were initiated with 3 $\mu$M aaRS and incubated at 37 °C. At time = 1, 5, 10, 20, 30, 60 and 90 minutes 10 $\mu$L was spotted onto Whattman paper and immediately submerged in 10 % (w/v) trichloroacetic acid (TCA). Papers were washed 3 x 15 mins in 10 % (w/v) TCA and 1 x 15 mins in 100 % ethanol, then left to air dry. To obtain a total count (TC) of radioactivity, 5 $\mu$L of a 1/10 dilution was spotted onto paper but not submerged in 10 % (w/v) TCA. All papers were added to 5 mL scintillant and counted.

### 2.12.2 Radioactive preparative charging assay

To test the purification procedure, and to generate charged tRNA for a trial MurM activity assay, a preparative charging experiment with AlaRS and cognate amino acid [$^3$H] L-Ala was conducted using *E. coli* tRNA.

Reactions containing 30 mM HEPES, 15 mM $MgCl_2$, 25 mM KCl, 5 mM DTT, 2 mM ATP (pH 7.5), 330 $\mu$g crude *E. coli*tRNA, 0.045 mM [$^3$H] L-Ala, 5 mM DTT were

incubated at 37 °C for 10 minutes. 6 $\mu$M AlaRS was added to initiate the reaction, mixed and incubated for 1 hour. 20 $\mu$L was spotted onto Whatmann paper and submerged in 10 % (w/v) TCA. Total count (TC) of radioactivity was measured by spotting 5 $\mu$L of a 1/10 dilution onto Whattman paper. 20 $\mu$L 3 M sodium acetate (pH 5.0) was added to the sample. 205 $\mu$L phenol was added, sample was vortexed and centrifuged at 13,000 *xg* for 3 mins. Aqueous phase was retained and 520 $\mu$L ice cold ethanol was added. Sample was incubated at -20 °C for 15 minutes to precipitate tRNA. tRNA was pelleted by centrifugation at 13,000 *xg* for 3 mins, supernatant was discarded and pellet was dried in a dessicator. Pellet was resuspended in 50 $\mu$L 3 mM sodium acetate (pH 5.0), 2 $\mu$L was spotted onto Whattman paper and submerged in 10 % (w/v) TCA. Papers were washed 3 x 15 mins in 10 % (w/v) TCA and 1 x 15 mins in 100 % ethanol, then left to air dry. All papers were added to 5 mL scintillant and counted. tRNA was stored at -80 °C.

### 2.12.3 Radiolabelled MurM activity assay

During these experiments, aminoacyl-tRNA were kept on ice to minimise deacylation.

A trial MurM activity assay was conducted using *E. coli* [$^3$H]alanyl-tRNA$^{Ala}$ prepared in Section 2.12.2 experiment. 0.15 mM Lipid II(Lys) or iGln Lipid II(Lys) was dried down and resuspended in 150 $\mu$L 50 mM HEPES, 10 mM $MgCl_2$, 30 mM KCl, 1.5 % (w/v) CHAPS (pH 7.6). Reactions comprising 50 mM HEPES, 10 mM $MgCl_2$, 30 mM KCl, 1.5 % (w/v) CHAPS (pH 7.6), 2 mM DTT, 1 mM L-Ala, 0.005 mg/ml $gMurM_{159}$ were initiated with 2.5 $\mu$M and 1.94 $\mu$M *E. coli* [$^3$H]alanyl-tRNA$^{Ala}$ for reactions containing Lipid II(Lys) and iGln Lipid II(Lys) respectively. Reactions were incubated at 37 °C for 1 hour. 1 x incubation volume of ice cold pyridinium acetate was added and sample was vortexed. 2 x incubation volume of N-butanol was added, sample was vortexed and centrifuged at 13,000 *xg* for 3 mins in a chilled centrifuge. The top phase was removed and added to 2 x incubation volume of $H_2O$ , sample was vortexed and centrifuged at 13,000 *xg* for 3 mins. The top phase was

added directly to scintillation vials and counted.

Negative controls were conducted with reactions in the absence of MurM, in the presence of RNase and in the absence of Lipid II(Lys).

## 2.13 Computational modelling of MurM

### 2.13.1 Homology modelling of MurM

*S. pneumoniae* gMurM$_{159}$, *S. aureus* FemX, *S. aureus* FemA (PDB ID: 1LRZ) and *W. viridescens* FemX (PDB ID: 3GKR) were aligned using EMBOSS Needle CLUSTAL (Madeira *et al.*, 2019) and sequence identity and similarity determined.

The structure of *S. aureus* FemX (unpublished work from Roper group) was used as the template for homology modelling due to its high relatedness with MurM. FemX and MurM sequences were aligned and using MODELLER (Eswar *et al.*, 2006; Martí-Renom *et al.*, 2000; Šali and Blundell, 1993; Fiser *et al.*, 2000) an initial test model was generated to verify the validity of the template and the alignment. We evaluated this model by computing its energy profile using the high-resolution version of the Discrete Optimized Protein Energy (DOPE-HR) (Shen and Sali, 2006), smoothed via window averaging with size 15 residues. The profiles of template and model were compared (Appendix 7.8), refinement was conducted in the region between Lys230 and Pro299. Thus, we proceeded to create 64 different models, refining their secondary structure using the slowest annealing protocol available in Modeller. The loop regions of each of the 64 base models were refined 16 independent times also using a very slow annealing procedure. Subsequently, we evaluated and ranked each of the resulting 1024 models using DOPE-HR as well as the Statistically Optimized Atomic Potentials (SOAP) (Dong *et al.*, 2013). The 10 best scoring models for each score were selected and eliminated based on the number of physical constraint violations present.

The final model of MurM (MurM159) was aligned with the previous MurM

model (Fiser *et al.*, 2003) or *W. viridescens* Femx homologues (Fonvielle *et al.*, 2013; Biarrotte-Sorin *et al.*, 2004) for visualisation and analysis in PyMOL (Version 2.1.0).

All sequences, except *S. pneumoniae* $MurM_{Pn16}$ and *S. pneumoniae* $MurM_{159}$ (Appendix 7.3), were obtained from the GenBank database (https://www.ncbi.nlm.nih.gov/genbank/). Pairwise sequence alignments were generated using Jalview (Version 2.10.5) (Waterhouse *et al.*, 2009). Ramachandran plots were generated using MolProbity (Chen *et al.*, 2010; Lovell *et al.*, 2003).

### 2.13.2 Molecular docking of truncated Lipid II(Lys) to MurM

Since the Lipid II(Lys) was too large for use in molecular docking, the hydrocarbon tail was shortened to create a smaller substrate (referred to as truncated Lipid II(Lys). The truncated Lipid II(Lys) was drawn in ChemDraw Professional (Version 17.1) and converted to a pdb file using Avogadro (Version 1.2.0). To prepare the ligand file for docking, the protonation state in $H_2O$ at pH 7.4 was computed. Subsequently the equilibrium geometry minimizing the potential energy was computed using the GAFF force field (Wang *et al.*, 2004) from within the Avogadro2 software (Hanwell *et al.*, 2012). Molecular docking was conducted using AutoDock Vina (Trott and Olson, 2010) and pdbqt files were generated (from the pdb files of receptor model and ligands) using AutoDock Tools (Morris *et al.*, 2009). Initially the location of the binding site was verified by providing the algorithm with a search space that included the whole protein, requesting the 10 best positions. The top 10 poses were all within the same pocket on the molecular surface of the protein, very strongly implying that the pocket is indeed the binding site. Docking was then repeated, restricting the search space to the binding pocket to obtain the final docked conformation.

### 2.13.3 Molecular dynamics of MurM interactions with the membrane

Molecular dynamics (MD) simulations of MurM with the membrane were conducted in collaboration with Prof. Syma Khalid (University of Southampton). Simulations and initial analysis were conducted by Jonathan Shearer (PhD student - University of Southampton).

#### Coarse-grained simulations

All coarse-grained simulations were carried out with the GROMACS package (Version 2018) and the Martini (Version 2.2) forcefield (Abraham *et al.*, 2015; de Jong *et al.*, 2012).

Initial protein coordinates from the MurM homology model were used to generate a coarse-grained model using the Martini script. The protein was coarse grained to the ElNeDyn model with an elastic network strength and cutoff of 500 kJmol-1nm-2, respectively.

The Lipid II(Lys) model for inclusion in the membrane was parameterised using a United atom model (Gromos 53a6) generated by the Automated topology builder (ATB) web-interface. Following this the coarse-grained mapping was decided iteratively and the bonded terms fitted with PyCGTOOl (Graham *et al.*, 2017).

Simulations were conducted with three different membrane systems. System 1 contained 75 % phosphatidylethanolamine (PE) and 25 % phosphatidylglycerol (PhG), system 2 contained 72 % PE, 16 % PG and 12 % cardiolipin (CL) and system 3 contained 72 % PE, 12 % PG and 16 % cardiolipin (CL). The membrane systems of $\sim$16x16x11.5 nm were generated with the Martini-maker module of the CHAMRMM-GUI web interface (Jo *et al.*, 2017). $Na^{2+}$ ions were then added to neutralise each system and equilibration was conducted using the usual CHARMM-GUI protocol. The temperature of all simulations was 313 K and was controlled with a stochastic velocity rescale thermostat with a coupling constant

of 1.0 ps. The CHARMM GUI protocol involved initial steepest descent energy minimisations, followed by a series of NPT equilibrations with 5, 10, 15 and 15 fs timesteps for a cumulative total of around 20 ns. During this procedure the Berendsen barostat was employed with a coupling constant of 4.0 ps (Berendsen *et al.*, 1984). Following equilibration ∼1 % of Lipid II(Lys) (10 in total) was added to each membrane. Minimisations and a 10 ns equilibration were conducted and then followed by a 2  production run to ensure sufficient mixing of all the lipid components. Note that all production runs were carried out using a 10 fs timestep and a Parrinello-Rahman semi-isotropic barostat with a 12 ps coupling constant (Parrinello and Rahman, 1981). The Lennard-Jones potential was cutoff using the Potential shift Verlet scheme at long ranges. The reaction field method was used for electrostatics calculations, with dielectric constants of 15 and infinity for charge screening in the short- and long-range regimes, respectively. The short-range cutoff for both non-bonded and electrostatic interactions was 1.2 nm. The size of the system was increased to ∼32 nm in dimension perpendicular to the membrane normal and the MurM was added in a random orientation around 8 nm above each membrane. Addition of $H_2O$ and 0.15 M NaCl was added to provide biologically relevant salt concentrations. Around 10 % of $dH_2O$ molecules were changed to antifreeze particles if systems froze during simulation. After an initial steepest descent minimisation and 1 ns equilibration, in which the protein backbone was restrained with 1000 $kJmol^{-1}.nm^{-2}$ harmonic restraints,  production runs were carried out with 6 repeats per system.

**All-atom simulations**

All atomistic simulations were conducted using the CHARMM36m 2018 forcefield (Huang *et al.*, 2017). The Lipid II(Lys) model used here was used in previous studies (Witzke *et al.*, 2016). Equilibration steps were for increasing timesteps of 1 and 2 fs were conducted with a berendsen thermostat and barostat with coupling constants of 1.0 and 5.0 ps, respectively. During production runs a timestep of 2 fs was used

and the pressure (1 atm) regulated with a semi-isotropic Parrinello-Rahman barostat with a coupling constant of 5.0 ps. The Lennard-Jones potential was cutoff with the Force-switch modifier from 1.0 to 1.2 nm. The short range cutoff for the electrostatic interaction was also 1.2 and the Particle mesh Ewald (PME) algorithm was used for the long-range regime. For each membrane system two distinct orientations for MurM relative to the membrane were chosen. These systems were then backmapped to the all-atoms model as previously described. After the reverse transformation was carried out, each system was cropped in the z dimension to a height of 16.5 nm, to remove the now unnecessary $H_2O$ molecules. Each system was minimised and equilibrated for a total of 1 ns, while the backbone of the protein was restrained with 1000 $kJmol^{-1}.nm^{-2}$ harmonic restraints. 2 production runs were carried out for each system for 300 ns.

### 2.13.4 Analysis of all-atom simulations

Analysis was carried out over the final 150 ns of each simulation, unless stated otherwise. All simulations were visualised using Visual Molecular Dynamics (VMD) or PyMOL (Version 2.2.0). Hydrogen bonding analysis was performed with the VMD hydrogen bonding extension. Other analysis tools were written with a combination of GROMACS tools and in house scripts that utilised the python module MDAnalysis (Gowers *et al.*, 2016). The depletion/enrichment (D-E) indices were determined by first counting the number of lipids with a center of geometry within 1.4 nm of the protein and then comparing this number to the number expected in the bulk of the membrane, using the procedure described by Corradi *et al.* (2018). The depletion/enrichment index was obtained by dividing the lipid composition in the 1.4 nm shell around the protein by the bulk membrane composition. Thus a D-E index $>1$ indicates enrichment, while a D-E index $<1$ indicates depletion. The D-E index was determined for the last 150 ns of each simulation in 50 ns blocks for both repeats. This gave 5 D-E index values for each lipid, from which the average and standard deviation was calculated. The enrichment maps were generated by

first determining the 2D density map of the membrane using the GROMACS tool densmap. Following this the enrichment was determined using the procedure described by Corradi *et al.* (2018).

### 2.13.5 Flexibility simulation and homo-/hetero-dimerisation of MurM

Pdb2movie (Jimenez-Roldan *et al.*, 2012; Römer *et al.*, 2016) was used to rapidly identify the potential movements and flexibility of the whole protein, by combining three software packages; ELNEMO, FIRST and FRODA. Default parameters were used with the following modifications; number of conformers, 5000; energy cutoff value, 1.0 and Modes 7, 8, 9 and 10.

# Chapter 3

# Effect of $H_2O_2$ on aminoacyl-tRNA synthetases

The authors would like to acknowledge the contribution of the WPH Proteomics Research Technology Platform (RTP), Gibbet Hill Road, University of Warwick, UK, for the peptide based mass spectrometric analysis of proteins.

## 3.1 Introduction

*S. pneumoniae* is a facultative anaerobe, in aerobic conditions pyruvate oxidase (SpxB) converts pyruvate into acetyl phosphate, which can be converted to acetate to generate an additional 2 molecules of ATP per molecule of glucose. However, this additional ATP comes at a cost; $H_2O_2$ is produced as a toxic biproduct of SpxB. Intriguingly, despite producing high levels of endogenous $H_2O_2$ *S. pneumoniae* does not possess the global regulators (OxyR or PerR), or either of the $H_2O_2$ scavenging enzymes (catalase and NADH peroxidase) that are present in many other species (Pericone *et al.*, 2003). Very little is known about how *S. pneumoniae* protects itself against oxidative stress caused by $H_2O_2$, but studies indicated that SpxB activity itself confers a protective effect (Pericone *et al.*, 2003).

The effects of oxidative stress on translation and cell viability have been well documented and high levels of $H_2O_2$ are known to increase mis-aminoacylation and translational error in other organisms (Bullwinkle *et al.*, 2014; Costa *et al.*, 2007; Ling and Söll, 2010). Ling and Söll (2010) showed previously that hydrogen peroxide oxidised cysteine 182 of *E. coli* ThrRS, resulting in a reduction of editing activity and an increase in seryl-$tRNA^{Thr}$ production. Both ThrRS and AlaRS are prone to mis-activation of serine and therefore posses pre- and post-transfer editing mechanisms to minimise mis-aminoacylation of $tRNA^{Thr}$ and $tRNA^{Ala}$ with L-Ser. Additionally, serine is incorporated to the first position of an indirect dipeptide cross-link in *S. pneumoniae* PG by MurM and unpublished data (Dr Adrian Lloyd) indicates that MurM has a preference for mis-aminoacylated tRNA. As such the focus of this chapter is to understand the effects of $H_2O_2$ on amino-acid activation, pre-transfer and post-transfer editing of ThrRS and AlaRS in *S. pneumoniae*.

In contrast to the work of Ling and Söll (2010), this chapter utilised a novel post-transfer editing assay developed by Dr Adrian Lloyd (unpublished) (Figure 2.3). Previous investigations by Ling and Söll (2010) relied on a radiolabelled ATP consumption end point assay (hereon in referred to as the end point assay). The end point assay measures the overall amount of editing by monitoring ATP consumption, ATP is consumed to produce an aa-AMP which is subsequently converted to aa-tRNA. It works on the premise that where editing is occurring and aa-AMP or aa-tRNA are hydrolysed back to aa and tRNA, multiple rounds of aminoacylation and hence ATP hydrolysis can occur per molecule of tRNA, whilst in the absence of editing (e.g. in the presence of a cognate amino acid) only one round of ATP consumption will occur per molecule of tRNA. It is therefore possible to follow the combined pre- and post-transfer editing activities of the aaRS using the end point assay. Since this is an end point assay, only the final ATP consumption after a specific time period is measured. In contrast, the assay used in this chapter is a novel continuous spectrophotometric assay, which generates a background rate in the absence of tRNA (pre-transfer editing rate) and a subsequent

rate following the addition of the tRNA (combination of pre- and post-transfer editing). This assay works by following the production of AMP which occurs during the hydrolysis of aa-AMP (pre-transfer editing), or transfer of the aminoacyl moiety from the adenylate to the tRNA where the aa-tRNA is hydrolysed releasing AMP (post-transfer editing). The production of AMP is coupled by myokinase, PK and LDH, to the conversion of NADH to $NAD^+$ which can be monitored by a loss in absorbance at 340 nm (Figure 2.3). This assay provides continuous data for kinetic analysis and also permits the distinction between pre-transfer editing and post-transfer editing mechanisms.

## 3.2 Aims

This chapter demonstrates the purification of *S. pneumoniae* AlaRS, SerRS and ThrRS. The catalytic efficiency of cognate and non-cognate amino acid activation by AlaRS and ThrRS was investigated. In addition the effect of $H_2O_2$ on amino acid activation, pre-transfer and post-transfer editing activities of *S. pneumoniae* AlaRS and ThrRS was determined and the site of oxidation investigated using mass-spectrometry.

## 3.3 Purification of aminoacyl-tRNA synthetases

AlaRS, SerRS and ThrRS were overexpressed, purified and quantified as per Section 2.5. Figure 3.1 a), b) and c) respectively, shows the purity of these proteins at each stage of the purification procedure.

For AlaRS a significant increase in purity was achieved by IMAC and a small increase in purity was achieved during the second, anion exchange chromatography step. SEC as an alternative or additional chromatography step, did not increase purity of the protein sample and so was eliminated to reduce the length of the purification procedure. A strong band corresponding to the expected mass of AlaRS at $\sim$97 kDa, was observed by SDS-PAGE. In addition there were a small number of

weaker bands which may represent contaminating proteins or products of AlaRS proteolysis (Figure 3.1a). The final purity of AlaRS, post anion exchange, was estimated to be >60 % using ImageJ (Version 1.51).

For SerRS a significant increase in purity was achieved by IMAC and SEC purification steps. A strong band corresponding to the expected mass of SerRS at ∼50 kDa was observed by SDS-PAGE, the final purity of SerRS was estimated to be ∼90 % using ImageJ (Version 1.51).

For ThrRS a significant increase in purity was achieved during both the IMAC and SEC steps. The final MonoQ chromatography step did not further enhance the sample's purity. A strong band corresponding to the expected mass of ThrRS at ∼75 kDa was observed, there are also some additional weaker bands at ∼55 kDa which may be contaminating proteins or degradation products. The final purity of ThrRS was estimated to be >75 % using ImageJ (Version 1.51).

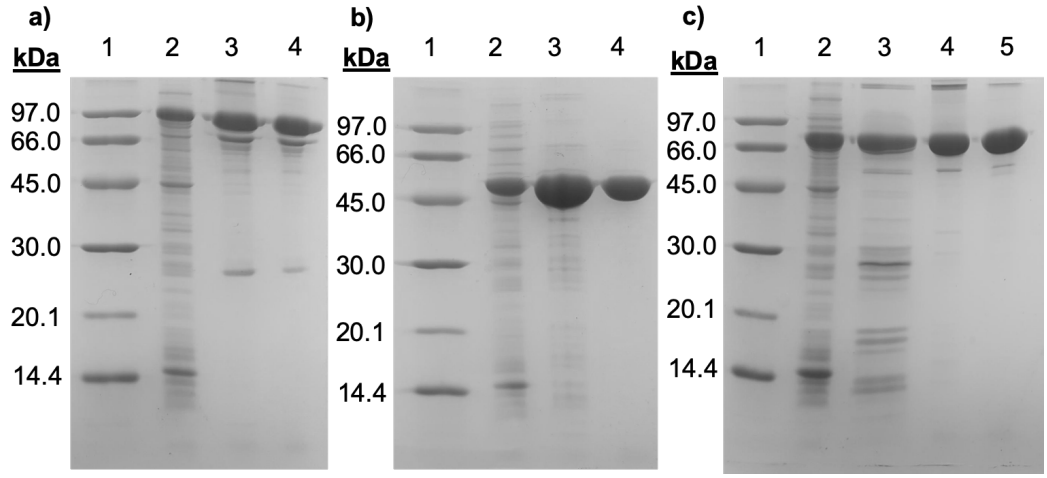



**Figure 3.1: 12 % SDS-PAGE showing protein purification summary.** a) Summary gel for the purification of AlaRS. Lane 1 - Amersham$^{TM}$ protein marker, Lane 2 - 10 µg 50,000 *xg* supernatant, Lane 3 - 10 µg pooled IMAC fractions, Lane 4 - 10 µg pooled anion exchange fractions. b) Summary gel for the purification of SerRS. Lane 1 - Amersham$^{TM}$ protein Marker, Lane 2 - 10 µg 50,000 *xg* supernatant, Lane 3 - 10 µg pooled IMAC fractions, Lane 4 - 10 µg pooled SEC fractions. c) Summary gel for the purification of ThrRS. Lane 1 - Amersham$^{TM}$ protein marker, Lane 2 - 10 µg 50,000 *xg* supernatant, Lane 3 - 10 µg pooled IMAC fractions, Lane 4 - 10 µg pooled SEC fractions, Lane 5 - 10 µg pooled anion exchange fractions.

## 3.4 Effect of $H_2O_2$ on amino acid activation

The first step of aminoacylation is activation of the amino acid by the aaRS. This is a condensation reaction between amino acid and ATP forming an aaRS bound aa-AMP. The consumption of ADPNP and concomitant pyrophosphate-dependent turnover of aa-AMP to form ATP is monitored using a coupled reaction which measures the production of NADPH and an associated increase in absorbance at 340 nm (see Figure 2.2).

### 3.4.1 Michaelis-Menten kinetics of amino acid activation by AlaRS with L-alanine and L-serine

Figure 3.2 shows the Michealis-Menten kinetics of *S. pneumoniae* AlaRS amino acid activation with cognate and non-cognate amino acids (L-Ala and L-Ser respectively). AlaRS has a catalytic efficiency ($k_{cat}{}^{App}/k_m{}^{App}$) of 73.4 $min^{-1}.mM^{-1}$ with L-Ala, however it is also capable of incorrectly activating L-Ser with a much lower $k_{cat}{}^{App}/k_m{}^{App}$ of 0.15 $min^{-1}.mM^{-1}$. Despite AlaRS having a selectivity for L-Ala over L-Ser of 489, it is clearly capable of mis-activation with L-Ser at a rate greater than that consistent with the error frequency of translation. This demonstrates the importance of both pre- and post-transfer editing mechanisms.

### 3.4.2 Michaelis-Menten kinetics of amino acid activation by ThrRS with L-threonine

Figure 3.3 shows the catalytic efficiency of ThrRS with L-Thr, which was determined to be 1.49 $mM^{-1}.min^{-1}$ is $\sim$ 50-fold lower than that of AlaRS. This reduced catalytic efficiency is due to a combination of comparatively low $k_{cat}{}^{App}$, indicating low turnover, and high $k_m{}^{App}$, suggesting low binding affinity, for *S. pneumoniae* ThrRS. In order to determine if an alternate ATP analogue such as adenosine 5′-($\beta$-$\gamma$-methylene) triphosphate (ADPCP) was a better substrate for ThrRS the assay

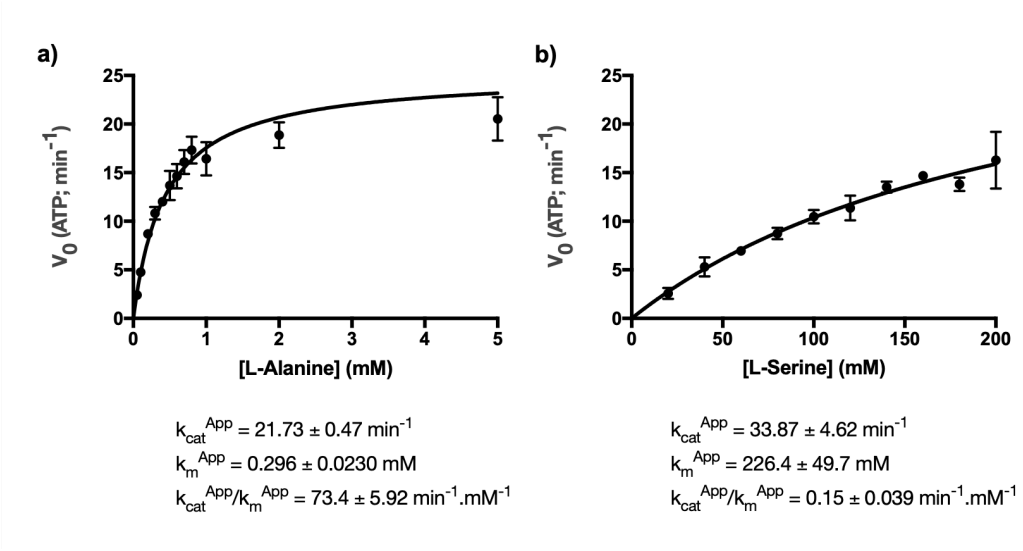



**Figure 3.2: Michaelis-Menten kinetics of AlaRS.** A plot of the reaction velocity ($V_0$) as a function of a) cognate amino acid (L-Ala) concentration and b) non-cognate amino acid (L-Ser) concentration. Apparent $k_{cat}$, $k_m$, and $k_{cat}/k_m$ are presented. $R^2$ = 0.96 and 0.94 for a) and b) respectively. Mean and standard deviation were plotted from triplicate data sets. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation.

was repeated (Figure 3.3 b)). Whilst, ADPCP did improve the turnover of ThrRS by 2.5-fold, only a very small change in binding affinity for L-Thr was observed. The $k_{cat}{}^{App}/k_m{}^{App}$ of ThrRS with L-Thr, was determined to be 4.49 $min^{-1}.mM^{-1}$ and, even with this improved catalytic activity it was still not possible to obtain kinetic data for mis-activation of ThrRS with L-Ser.

### 3.4.3 Amino acid activation assay limitations

The purpose of this assay was to demonstrate that AlaRS and ThrRS can mis-activate with serine. This was only possible for AlaRS since the kinetics of ThrRS were so poor that it was not possible to monitor mis-activation of serine using this assay. Traditionally, amino acid activation studies have been conducted using the pyrophosphate exchange assay (Francklyn *et al.*), however in this study a coupled spectrophotometric assay, referred to as the ADPNP assay, was employed (Figure 2.2). The pyrophosphate exchange assay measures the incorporation of pyrophosphate into ATP through the intermediacy of the enzyme.AMP-AA complex

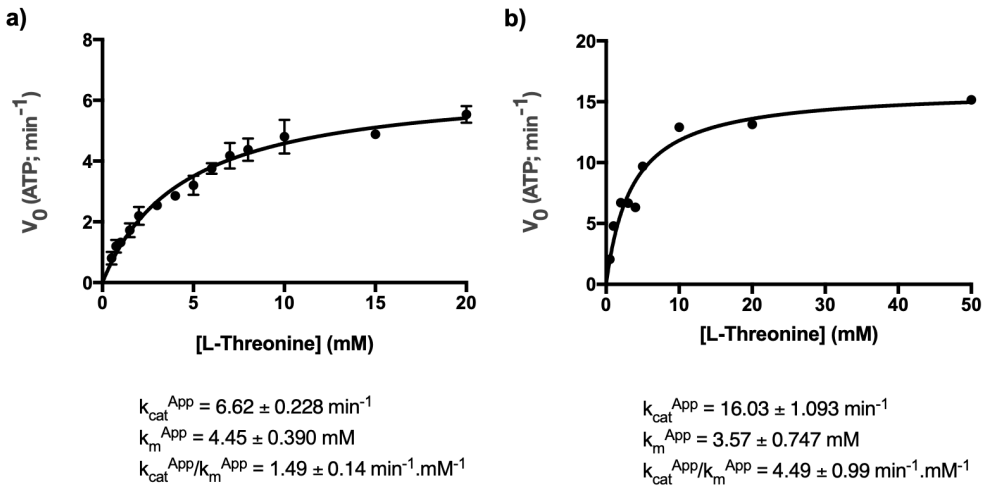



**Figure 3.3: Michaelis-Menten kinetics of ThrRS.** A plot of the reaction velocity ($V_0$) as a function of cognate amino acid (L-Thr) concentration. a) Assay using ADPNP as the ATP analogue. Data collected in triplicate with average and standard deviation plotted. $R^2$ = 0.96. b) Assay using ADPCP as the ATP analogue. Data collected in singlicate. $R^2$ = 0.94. Apparent $K_{cat}$, $K_m$, and catalytic efficiency $K_{cat}/K_m$ are presented. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation.

(Francklyn *et al.*). As with all experimental procedures both assays have their advantages and disadvantages, the ADPNP assay eliminates the necessity to work with radiolabelled compounds and is a continuous assay, but utilises an ATP analogue (ADPNP or ADPCP) as substrate, whilst the pyrophosphate exchange assay uses the natural ATP substrate but requires radiolabelled pyrophosphate and is a non-continuous assay. Amino acid activation of L-Thr by *E. coli* ThrRS, using the [$^{32}$P]-pyrophosphatase assay, was reported to be 8400 $min^{-1}.mM^{-1}$ (>5000-fold higher than reported here) Ling and Söll (2010). The reduced catalytic activity of AlaRS and ThrRS reported here may indicate that aaRS's have a particularly high selectivity for ATP over analogues, and so the pyrophosphatase assay may be more suitable for characterising the kinetic parameters of these enzymes.

### 3.4.4 Effect of $H_2O_2$ on amino acid activation

ADPNP assays for AlaRS and ThrRS were conducted at a saturating concentration of cognate and non-cognate amino acid (as determined previously), and the effect of 50

mM $H_2O_2$ on amino acid activation was measured. Unless otherwise stated data was collected in triplicate (n=3) and analysed using a paired t-test or a one-way analysis of variance (ANOVA) (for multiple comparisons) with a 95 % confidence interval ($p \leq 0.05$). $H_2O_2$ had no statistically significant effect on amino acid activation of cognate or non-cognate amino acids by AlaRS or ThrRS (Figure 3.4). Furthermore, this demonstrates that $H_2O_2$ had no impact on the enzymatic coupling system.

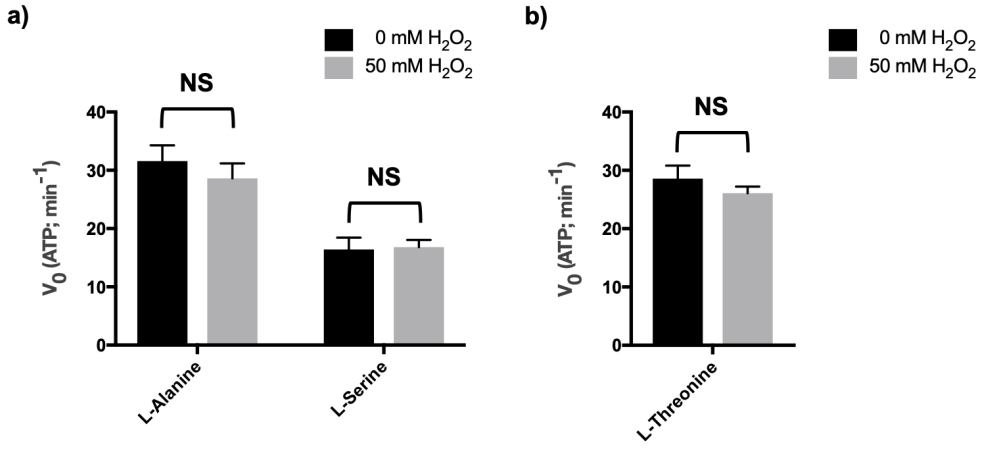



**Figure 3.4: Effect of $H_2O_2$ on aminoacylation activity of AlaRS and ThrRS.** a) Initial velocity of AlaRS amino acid activation with L-Ala or L-Ser in the presence and absence of 50 mM $H_2O_2$. b) Initial velocity of ThrRS amino acid activation with L-Thr in the presence and absence of 50 mM $H_2O_2$. Mean and standard deviation were plotted from triplicate data sets. NS = not statistically significant as determined by a paired t-test or one-way ANOVA. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation.

Amino acid activation of L-Ala by AlaRS was not statistically different ($p > 0.05$) between reactions containing 0 mM $H_2O_2$ ($31.57 \pm 1.568$ $min^{-1}$) and 50 mM $H_2O_2$ ($28.62 \pm 1.475$ $min^{-1}$). Amino acid activation of L-Ser by AlaRS was not statistically different ($p > 0.05$) between reactions containing 0 mM $H_2O_2$ ($16.43 \pm 1.171$ $min^{-1}$) and 50 mM $H_2O_2$ ($16.81 \pm 0.7082$ $min^{-1}$). Amino acid activation of L-Thr by ThrRS was not statistically different ($p > 0.05$) between reactions containing 0 mM $H_2O_2$ ($28.58 \pm 1.298$ $min^{-1}$) and 50 mM $H_2O_2$ ($26.09 \pm 0.6438$ $min^{-1}$).

## 3.5 Effect of $H_2O_2$ on pre- and post-transfer editing of AlaRS and ThrRS

Figures 3.5 and 3.6 show the full data sets of pre- and post-transfer editing rates plotted against $H_2O_2$ concentration for AlaRS and ThrRS respectively. In addition a bar chart corresponding to each data set summarises the results by comparing data at both the lowest and highest concentrations of $H_2O_2$. For both AlaRS and ThrRS increasing $H_2O_2$ concentrations positively correlated with loss of post-transfer editing activity but had no effect on pre-transfer editing activity. AlaRS and ThrRS in the presence of cognate amino acids (L-Ala and L-Thr respectively) were used as negative controls, to show the absence of pre- or post-transfer editing activity.

0.5 $\mu$M AlaRS required $\sim$75 mM $H_2O_2$ (150,000 x molar concentration) to reduce the rate of post-transfer editing to background, whilst 10 $\mu$M ThrRS required $\sim$17.5 mM (1750 x molar concentration) to achieve background post-transfer editing activity. ThrRS is therefore $\sim$85 x more sensitive to $H_2O_2$ than AlaRS.

The novel editing assay was used to investigate the effects of $H_2O_2$ on pre- and post-transfer editing of AlaRS. Unless otherwise stated data was collected in triplicate (n=3) and analysed using a paired t-test or a one-way ANOVA (for multiple comparisons) with a 95 % confidence interval ($p \leq 0.05$).

### 3.5.1 Effect of $H_2O_2$ on pre- transfer editing in AlaRS

Figure 3.5a and 3.5b show that pre-transfer editing of AlaRS occurs with L-Ser but not L-Ala (negative control), and that $H_2O_2$ has no effect on pre-transfer editing.

There is no statistically significant difference ($p > 0.05$) in pre-transfer editing rate of AlaRS with L-Ala between reactions containing 0 mM ($0.0024 \pm 0.0029$ min$^{-1}$) and 75 mM ($0.0024 \pm 0.00014$ min$^{-1}$, n=2) $H_2O_2$. At 0 mM $H_2O_2$ the rate of pre-transfer editing of AlaRS with L-Ala ($0.0024 \pm 0.0029$ min$^{-1}$) was statistically significantly lower than with L-Ser ($0.063 \pm 0.0062$ min$^{-1}$) (p=0.0015). This demonstrates that, as

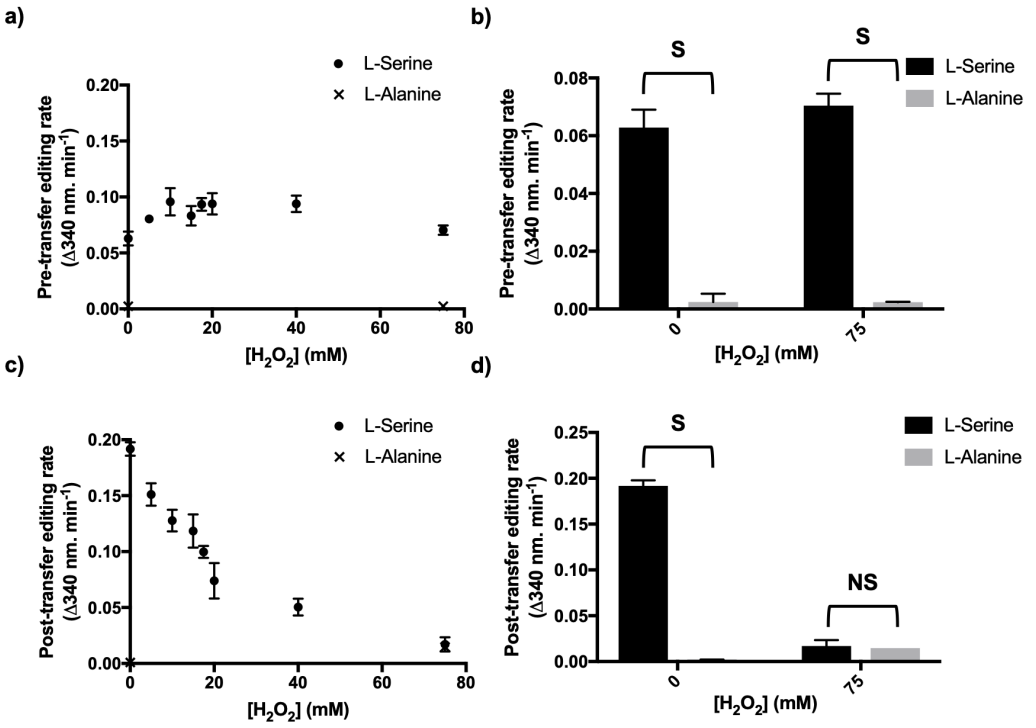



**Figure 3.5: Effect of $H_2O_2$ on pre-transfer and post-transfer editing rates of AlaRS.** 0.5 $\mu$M AlaRS with 200 mM L-Ala or 200 mM L-Ser were used in these spectrophotometric assays to obtain suitable rates of reaction. a) The rate of pre-transfer editing of L-Ser (or L-Ala as a negative control) by AlaRS at varying concentrations of $H_2O_2$. b) Bar chart showing pre-transfer editing rates of AlaRS at 0 and 75 mM $H_2O_2$ with both L-Ser and L-Ala (negative control). c) The rate of post-transfer editing of L-Ser (or L-Ala as a negative control) by AlaRS at varying concentrations of $H_2O_2$. d) Bar chart showing post-transfer editing rates of AlaRS at 0 and 75 mM $H_2O_2$ with both L-Ser and L-Ala (negative control). Mean and standard deviation were plotted from triplicate data sets. NS = not statistically significant, S = statistically significant, as determined by a paired t-test or one-way ANOVA. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation.

expected the cognate amino acid (L-Ala) undergoes very little pre-transfer editing when compared to non-cognate amino acid (L-Ser), and that $H_2O_2$ has no impact on the negative control.

In addition there was no statistically significant difference ($p>0.05$) in pre-transfer editing rate of AlaRS with L-Ser between reactions containing 0 mM ($0.063 \pm 0.0062$ $min^{-1}$) and 75 mM ($0.070 \pm 0.0041$ $min^{-1}$) $H_2O_2$. This demonstrates that increasing concentrations of $H_2O_2$ has no effect on the pre-transfer editing activity of AlaRS, and demonstrates that $H_2O_2$ had no impact on the enzymatic coupling system.

### 3.5.2 Effect of $H_2O_2$ on post-transfer editing in AlaRS

Figure 3.5c and 3.5d show that increasing concentrations of $H_2O_2$ positively correlated with the reduction in post-transfer editing activity of AlaRS. Post-transfer editing of AlaRS occured with L-Ser but not L-Ala (negative control) and high enough concentrations of $H_2O_2$ completely abolished post-transfer editing of AlaRS, reducing rates to those of the negative control.

There is no statistically significant difference ($p>0.05$) in post-transfer editing rate of AlaRS with L-Ala between reactions containing 0 mM ($0.00095 \pm 0.0012$ $min^{-1}$) and 75 mM ($0.015 \pm 0.000084$ $min^{-1}$, n=2) $H_2O_2$. Additionally at 0 mM $H_2O_2$ there was a statistically significant difference between the rate of post-transfer editing by AlaRS with L-Ala ($0.00095 \pm 0.0012$ $min^{-1}$) and L-Ser ($0.19 \pm 0.0060$ $min^{-1}$) ($p<0.0001$). This shows that as expected the cognate amino acid (L-Ala) undergoes very little post-transfer editing when compared to the non-cognate amino acid (L-Ser). The presence of $H_2O_2$ has no impact on the negative control.

The rate of post-transfer editing of AlaRS with L-Ser was statistically significantly higher at 0 mM $H_2O_2$ ($0.19 \pm 0.0060$ $min^{-1}$) than with 75 mM $H_2O_2$ ($0.017 \pm 0.0064$ $min^{-1}$) ($p<0.0001$). In addition there was no statistically significant difference between AlaRS with L-Ser at 75 mM ($0.017 \pm 0.0064$ $min^{-1}$) and AlaRS with L-Ala at 75 mM ($0.015 \pm 0.0085$ $min^{-1}$) ($p>0.05$) demonstrating that increasing concentrations of $H_2O_2$ reduces the post-transfer editing rate to that of the negative control.

### 3.5.3 Effect of $H_2O_2$ on pre- transfer editing in ThrRS

Figure 3.6a and 3.6b show that pre-transfer editing of ThrRS occurred with L-Ser but not L-Thr (negative control), and that $H_2O_2$ has no effect on pre-transfer editing.

There was no statistically significant difference ($p>0.05$) in pre-transfer editing rate of ThrRS with L-Thr between reactions containing 0 mM ($0.010 \pm 0.014$ $min^{-1}$) and 50 mM ($0.0.019 \pm 0.014$ $min^{-1}$, n=2) $H_2O_2$. In addition, there was no statistically

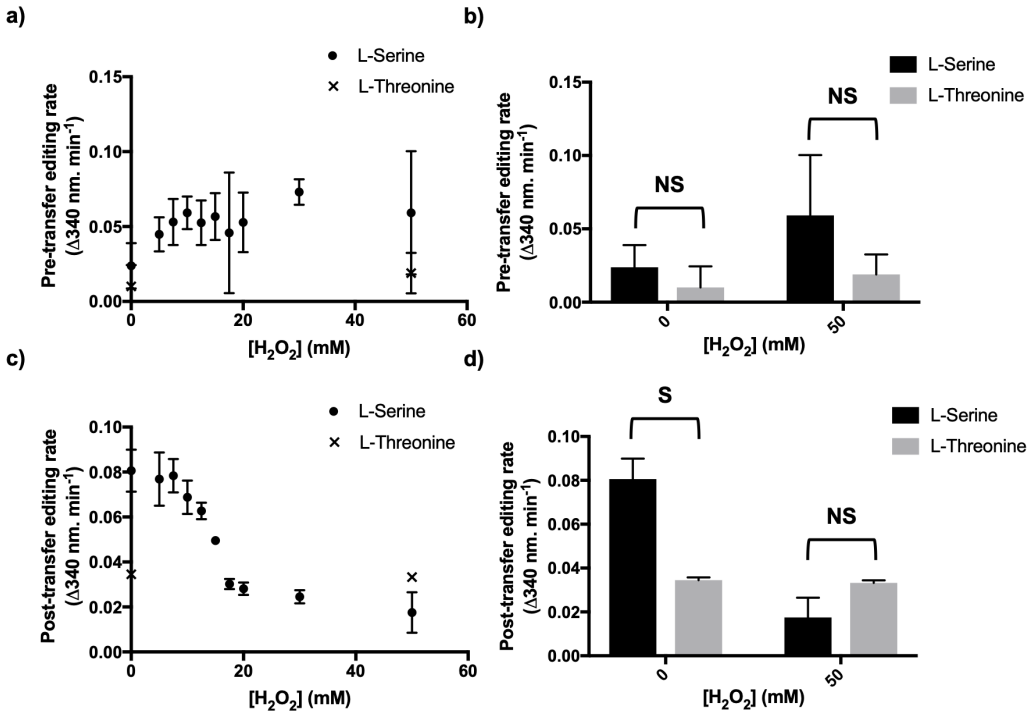



**Figure 3.6: Effect of $H_2O_2$ on pre-transfer and post-transfer editing rates of ThrRS.** 10 $\mu$M ThrRS with 50 mM L-Thr or 50 mM L-Ser was used in these spectrophotometric assays to obtain suitable rates of reaction. a) The rate of pre-transfer editing of L-Ser (or L-Thr as a negative control) by ThrRS at varying concentrations of $H_2O_2$. b) Bar chart showing pre-transfer editing rates of ThrRS at 0 and 50 mM $H_2O_2$ with both L-Ser and L-Thr (negative control). c) The rate of post-transfer editing of L-Ser (or L-Thr as a negative control) by ThrRS at varying concentrations of $H_2O_2$. d) Bar chart showing post-transfer editing rates of ThrRS at 0 and 50 mM $H_2O_2$ with both L-Ser and L-Thr (negative control). Mean and standard deviation were plotted from triplicate data sets. NS = not statistically significant, S = statistically significant, as determined by a paired t-test or one-way ANOVA. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation.

significant difference ($p>0.05$) between the rate of pre-transfer editing of ThrRS with L-Thr ($0.010 \pm 0.014$ min$^{-1}$) at 0 mM $H_2O_2$ and ThrRS with L-Ser ($0.024 \pm 0.015$ min$^{-1}$) . Taken together these results demonstrate that the rate of pre-transfer editing activity of ThrRS is the same in the presence of cognate and non-cognate amino acids, and that $H_2O_2$ has no impact on the negative control.

In addition, there was no statistically significant difference ($p>0.05$) in pre-transfer editing rate of ThrRS with L-Ser between reactions containing 0 mM ($0.024 \pm 0.015$ min$^{-1}$) $H_2O_2$ and 50 mM ($0.059 \pm 0.041$ min$^{-1}$) $H_2O_2$. This demonstrates that increasing concentrations of $H_2O_2$ has no effect on the pre-transfer editing activity

of ThrRS, and demonstrates that $H_2O_2$ had no impact on the enzymatic coupling system.

### 3.5.4 Effect of $H_2O_2$ on post-transfer editing in ThrRS

Figure 3.6c and 3.6d show that increasing concentrations of $H_2O_2$ positively correlated with the reduction in post-transfer editing activity of ThrRS. Post-transfer editing of ThrRS occurs with L-Ser but not L-Thr (negative control) and high enough concentrations of $H_2O_2$ completely abolish post-transfer editing of ThrRS, reducing rates to those of the negative control.

There was no statistically significant difference ($p>0.05$) in post-transfer editing rate of ThrRS with L-Thr between reactions containing 0 mM ($0.035 \pm 0.0013$ $min^{-1}$) and 50 mM ($0.033 \pm 0.0012$ $min^{-1}$, n=2) $H_2O_2$. Additionally at 0 mM $H_2O_2$ there was a statistically significant difference between the rate of post-transfer editing by ThrRS with L-Thr ($0.035 \pm 0.0013$ $min^{-1}$) and L-Ser ($0.081 \pm 0.0093$ $min^{-1}$) ($p=0.0004$). This shows that as expected the cognate amino acid (L-Thr) undergoes very little post-transfer editing when compared to the non-cognate amino acid (L-Ser), and that the presence of $H_2O_2$ had no impact on the negative control.

The rate of post-transfer editing of ThrRS with L-Ser was statistically significantly higher at 0 mM $H_2O_2$ ($0.081 \pm 0.0093$ $min^{-1}$) than with 50 mM $H_2O_2$ ($0.018 \pm 0.0091$ $min^{-1}$) ($p<0.0001$). In addition there was no statistically significant difference between ThrRS with L-Ser at 50 mM ($0.018 \pm 0.0091$ $min^{-1}$) and ThrRS with L-Thr at 50 mM ($0.033 \pm 0.0012$ $min^{-1}$, n=2) ($p>0.05$) demonstrating that increasing concentrations of $H_2O_2$ reduced the post-transfer editing rate to that of the negative control.

## 3.6 Effect of aaRS and $H_2O_2$ incubation time on post-transfer editing of L-Ser by AlaRS and ThrRS

The observation that increasing concentrations of $H_2O_2$ positively correlated with loss of post-transfer editing of aaRSs gave rise to two additional questions; was this directly due to oxidation of the aaRS by $H_2O_2$ alone and, if so, was this oxidation reaction a time dependent process. In order to investigate these, aaRS was incubated with $H_2O_2$ and at various time points the post-transfer editing activity was assayed as before. Since the oxidation of aaRS by $H_2O_2$ is a bimolecular reaction, the rate of this reaction will be dependant upon the relative concentration of both reactants. The molar ratio of aaRS:$H_2O_2$ from previous experiments (Section 3.5) which generated little/no reduction in post-transfer editing was used in these assays. The molar ratios of aaRS:$H_2O_2$ were 1:10,000 and 1:200 (equivalent to 5 mM $H_2O_2$ and 2 mM) for AlaRS and ThrRS respectively.

Figure 3.7 and 3.8 show that compared to 0 mM $H_2O_2$ control, incubation time has no effect on pre-transfer editing of serine. However, loss of post-transfer editing activity of mis-serylated $tRNA^{Ala}$ or $tRNA^{Thr}$ was positively correlated with incubation time in both AlaRS and ThrRS. Incubation of aaRS at room temperature without $H_2O_2$ was used as a negative control at the first and last time point, no effect on pre- or post-transfer editing was observed.

Unless otherwise stated data was collected in triplicate (n=3) and analysed using a paired t-test or a one-way ANOVA (for multiple comparisons) with a 95 % confidence interval ($p \leq 0.05$).

### 3.6.1 Effect of incubation time on pre- and post-transfer editing activity of AlaRS with L-Ser

At a fixed concentration of $H_2O_2$, the length of incubation time of AlaRS with $H_2O_2$ was shown to positively correlate to the loss of post-transfer editing. No effect on

pre-transfer editing activity of AlaRS was observed with varying incubation times.

Figure 3.7a shows that there was no difference in the rate of pre-transfer editing with respect to increased incubation time of AlaRS with 5 mM $H_2O_2$. In addition the negative controls demonstrated that incubation in the absence of $H_2O_2$ had no effect on pre-transfer editing of AlaRS; there was no statistically significant difference in the rate of pre-transfer editing of AlaRS between an incubation period of 0 mins ($0.036 \pm 0.0093$ min$^{-1}$) and 30 mins ($0.045 \pm 0.018$ min$^{-1}$) with $H_2O_2$. Additionally, there was no difference between pre-transfer editing of AlaRS with L-Ser after 0 mins incubation with ($0.036 \pm 0.0093$ min$^{-1}$) or without ($0.022 \pm 0.0053$ min$^{-1}$) $H_2O_2$. These results suggest that neither $H_2O_2$ or incubation time affect the rate of pre-transfer editing activity of AlaRS.

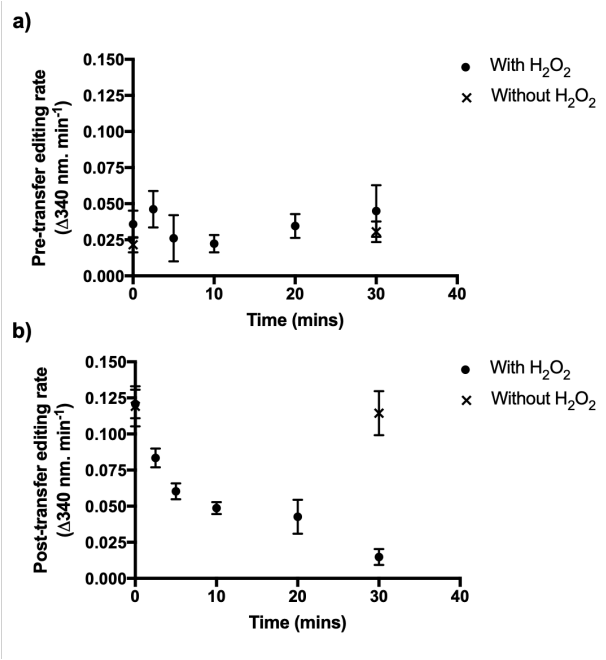



**Figure 3.7: Effect of pre-incubation time for AlaRS and $H_2O_2$ on pre- and post-transfer editing activity of L-Ser.** a) 30 minute time-course showing pre-transfer editing of AlaRS following incubation with $H_2O_2$. b) 30 minute time-course showing post-transfer editing of ThrRS following incubation with $H_2O_2$. The final concentration of AlaRS and $H_2O_2$ in the assay was 0.5 $\mu$M and 5 mM respectively. Incubation of AlaRS with 0 mM $H_2O_2$ was used as a negative control. Mean and standard deviation were plotted from triplicate data sets. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation.

Figure 3.7b shows that increasing incubation time of AlaRS with a fixed concentration of $H_2O_2$ positively correlated with a loss of post-transfer editing

activity. In the presence of $H_2O_2$ there was a statistically significant different reduction in post-transfer editing of AlaRS with L-Ser, between an incubation time of 0 mins ($0.12 \pm 0.010$ $min^{-1}$) and 30 mins ($0.015 \pm 0.0055$ $min^{-1}$) ($p<0.0001$ ). In contrast the negative control showed that there was no statistically significant difference ($p>0.05$) between the rate of post-transfer editing of AlaRS after an incubation time of 0 mins ($0.036 \pm 0.0093$ $min^{-1}$) or 30 mins ($0.045 \pm 0.018$ $min^{-1}$) in the absence of $H_2O_2$. These results demonstrate that the length of incubation time of AlaRS with $H_2O_2$ affects the editing activity of AlaRS with L-Ser. The negative control showed that there was no difference in post-transfer editing activity of AlaRS in the absence of $H_2O_2$, indicating that the presence of $H_2O_2$ does not have any adverse effect on other components of the assay (e.g. tRNA or enzymes from the coupling reaction).

### 3.6.2 Effect of incubation time on pre- and post-transfer editing activity of ThrRS with L-Ser

At a fixed concentration of $H_2O_2$, the length of incubation time of ThrRS with $H_2O_2$ was shown to positively correlate to the loss of post-transfer editing. No effect on pre-transfer editing activity of AlaRS was observed with varying incubation times.

Figure 3.8a shows that there was no difference in the rate of pre-transfer editing with respect to increased incubation time of ThrRS with 2 mM $H_2O_2$. In addition the negative controls demonstrated that incubation in the absence of $H_2O_2$ had no effect on pre-transfer editing of ThrRS. There was no statistically significant difference ($p>0.05$) on the rate of pre-transfer editing of ThrRS between an incubation period of 0 mins ($0.041 \pm 0.0034$ $min^{-1}$) and 40 mins ($0.034 \pm 0.0030$ $min^{-1}$) with $H_2O_2$. Additionally, there was no statistically significant difference ($p>0.05$) between pre-transfer editing of ThrRS after 0 mins incubation with ($0.041 \pm 0.0034$ $min^{-1}$) or without ($0.038 \pm 0.012$ $min^{-1}$) $H_2O_2$. These results suggest that neither $H_2O_2$ or incubation time affect the rate of pre-transfer editing activity of ThrRS.

Figure 3.8b shows that increasing incubation time of ThrRS with a fixed concentration of $H_2O_2$ positively correlated with a loss of post-transfer editing activity. In the presence of $H_2O_2$ there was a statistically significant different reduction in post-transfer editing activity of ThrRS, between an incubation time of 0 mins ($0.18 \pm 0.032$ $min^{-1}$) and 40 mins ($0.049 \pm 0.011$ $min^{-1}$) ($p=0.0003$). In contrast, the negative control showed that there was no statistically significant difference ($p>0.05$) between the rate of post-transfer editing of ThrRS after an incubation time of 0 mins ($0.18 \pm 0.014$ $min^{-1}$) or 40 mins ($0.14 \pm 0.023$ $min^{-1}$) in the absence of $H_2O_2$. These results demonstrate that the length of incubation time of ThrRS with $H_2O_2$ affects the editing activity of ThrRS with L-Ser. The negative control showed that there was no difference in post-transfer editing activity of ThrRS in the absence of $H_2O_2$, indicating that the presence of $H_2O_2$ does not have any adverse effect on other components of the assay (e.g. tRNA or enzymes from the coupling reaction).

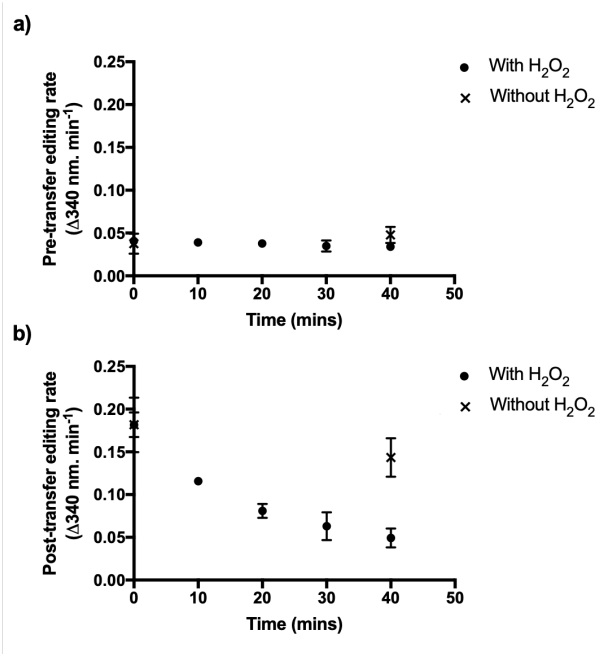



**Figure 3.8: Effect of pre-incubation time for ThrRS and $H_2O_2$ on pre- and post-transfer editing activity of L-Ser.** a) 40 minute time-course showing pre-transfer editing of ThrRS following incubation with $H_2O_2$. b) 40 minute time-course showing post-transfer editing of ThrRS following incubation with $H_2O_2$. The final concentrations of ThrRS and $H_2O_2$ in the assay were 10 $\mu$M and 2 mM respectively. Incubation of ThrRS with 0 mM $H_2O_2$ was used as a negative control. Mean and standard deviation were plotted from triplicate data sets. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation.

### 3.6.3 Additional experimental controls

In order to determine if the presence of DTT in the reaction mix was diminishing the observed effects of $H_2O_2$ on post-transfer editing rates, control experiments were conducted in the absence of DTT at incubation times of 10 mins for aaRSs (Figure 3.9a and b, for AlaRS and ThrRS respectively).

Unless otherwise stated data was collected in triplicate (n=3) and analysed using a paired t-test or a one-way ANOVA (for multiple comparisons) with a 95 % confidence interval ($p \leq 0.05$).

There was no statistically significant difference ($p > 0.05$) between assays conducted in the presence ($0.13 \pm 0.015$) or absence of DTT ($0.15 \pm 0.018$) when post-transfer editing of AlaRS with L-Ser was monitored following a 10 min incubation with 0 mM $H_2O_2$. Similarly, there was no statistically significant difference ($p > 0.05$) between assays conducted in the presence ($0.030 \pm 0.0036$) or absence of DTT ($0.035 \pm 0.0079$) when post-transfer editing of AlaRS with L-Ser was monitored following a 10 min incubation with 5 mM $H_2O_2$.

There was no statistically significant difference ($p > 0.05$) between assays conducted in the presence ($0.13 \pm 0.015$) or absence of DTT ($0.15 \pm 0.020$) when post-transfer editing of ThrRS with L-Ser was monitored following a 10 min incubation with 0 mM $H_2O_2$. Similarly, there was no statistically significant difference ($p > 0.05$) between assays conducted in the presence ($0.10 \pm 0.0034$) or absence of DTT ($0.11 \pm 0.0043$) when post-transfer editing of ThrRS with L-Ser was monitored following a 10 min incubation with 2 mM $H_2O_2$.

These results demonstrate that the presence of DTT in the assay mixture does not affect the post-transfer editing rate of either AlaRS or ThrRS.

In addition, to ensure that the measured effects of $H_2O_2$ were directly due to modified post-transfer editing activity, the effect of $H_2O_2$ on the coupled reaction was measured (Figure 3.9c and d). In this assay the rate limiting step is the

post-transfer editing mechanism and so the rate of the coupled system (comprising myokinase, PK and LDH) must be much faster. The coupled reaction was initiated by addition of AMP and can be seen to reach completion in $< 30$ s, this is unaffected by the addition of 50 mM $H_2O_2$. However, whilst this has no affect on the initial rate of reaction measured, interestingly the overall absorbance change of the reaction decreases in the presence of $H_2O_2$. This experiment shows that the rate of the coupled reactions are so rapid, even in the presence of $H_2O_2$, that any effect $H_2O_2$ may have on these enzymes would be negligible in this assay.

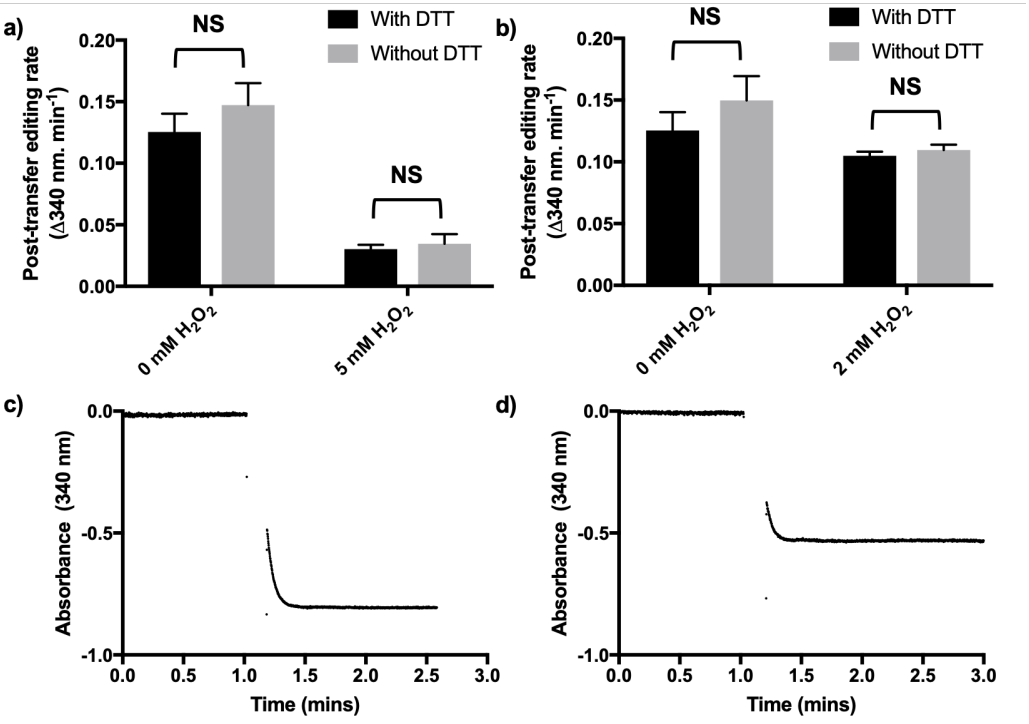



**Figure 3.9: Effect of $H_2O_2$ controls.** Effect of presence of DTT in master mix on Post-transfer editing of a) AlaRS and b) ThrRS, following 0 min and 10 min incubation with 50 mM $H_2O_2$. Spectrophotometric data of the coupled reactions occurring in real-time as detected by changes in absorbance at 340 nm c) in the absence of $H_2O_2$ d) in the presence of 50 mM $H_2O_2$. NS = not statistically significant, as determined by a paired t-test or one-way ANOVA. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation.

## 3.7 An alternate ThrRS construct

In addition to the low catalytic efficiency of ThrRS compared to AlaRS observed in the amino acid activation assay, the post-transfer editing assay required a 20-fold higher concentration of ThrRS than AlaRS in order to obtain a measurable editing rate. The ThrRS used by Ling and Söll (2010) and Wu *et al.* (2014) contained an N-terminal 6-His tag whilst the *S. pneumoniae* ThrRS used here possessed a C-terminal 6-His tag. If the C-terminal end of *S. pneumoniae* ThrRS is important for activity, the location of the 6-His tag may explain the marked reduction in catalytic efficiency of this enzyme. The pET22::ThrRS(spr1472) construct does not have a cleavage site present for removal of the C-terminal 6-His tag, and so in order to understand the importance of the C-terminus for activity, ThrRS was cloned into pET28(a) to engineer an N-terminally tagged construct (ThrRS(N-term)). ThrRS(N-term) was overexpressed, purified and quantified as outlined in Section 2.5.

For ThrRS(N-term) a significant increase in purity was achieved by IMAC and very little additional purity was achieved during the second chromatography step using SEC. The final purity of ThrRS(N-term) was estimated to be $\sim$ 63 % using ImageJ (Version 1.51).

ThrRS(N-term) was re-assayed using both the ADPNP amino acid activation assay, and the post-transfer editing assay. The activity of ThrRS(N-term) remained low and no significant improvement was achieved in either assay, therefore all subsequent assays were conducted with the C-terminally tagged ThrRS for consistency. This result suggests that the presence of a C-terminal his-tag does not explain the low activity of ThrRS when compared to AlaRS.

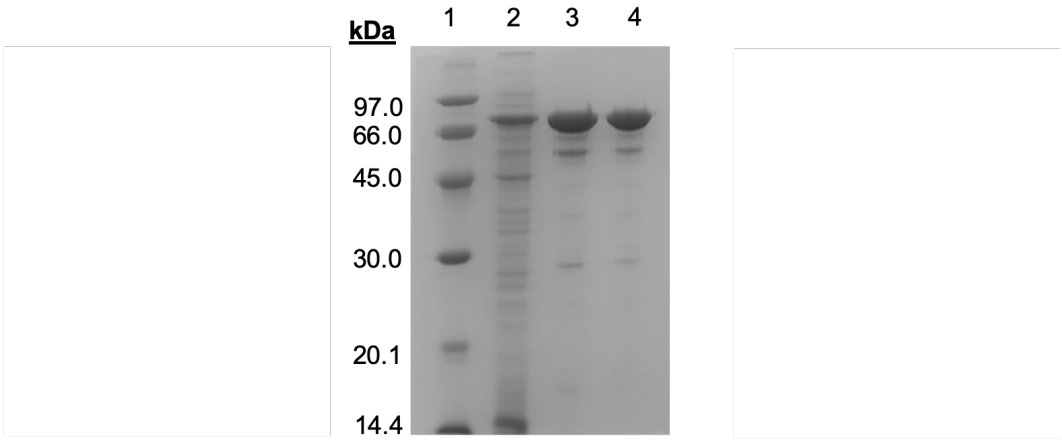



**Figure 3.10: 12 % SDS-PAGE showing ThrRS protein purification summary.** a) Summary gel for the purification of N-terminally tagged ThrRS. Lane 1 - Amersham Protein Marker, Lane 2 - 10 µg 50,000 x g supernatant, Lane 3 - 10 µg pooled IMAC fractions, Lane 4 - 10 µg pooled SEC fractions.

## 3.8 Mass Spectrometry to identify the modification made to ThrRS and AlaRS by $H_2O_2$

This section aims to demonstrate that AlaRS and ThrRS undergo stable oxidative modifications when treated with $H_2O_2$ and secondly to characterise which residues are modified (and to what oxidation state) in the conditions previously shown to abolish post-transfer editing activity. Cysteine and methionine both contain sulphur and so are the most likely amino acids to undergo oxidation (Ahmad *et al.*, 2017). Each residue can form multiple oxidation states (Table 3.1); cysteine residues can be oxidised upto 3 times creating sulphenic, sulphinic and sulphonic acids respectively, whilst methionine residues can exist in one of two oxidation states; methionine sulphoxide or methionine sulphone. The molecular weight of oxygen is 15.9949 Da, therefore it was anticipated that intact protein mass spectrometry would identify shifts in protein mass occurring as multiples of 15.9949 Da. *S. pneumoniae* AlaRS contains a single cysteine residue, whilst ThrRS contains two cysteine residues which are suitable targets for oxidation by $H_2O_2$ (Appendix 7.3).

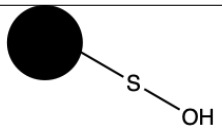
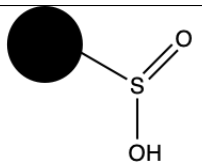
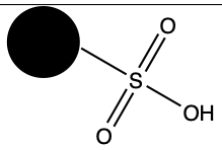
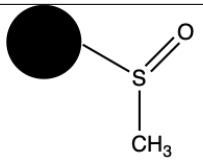
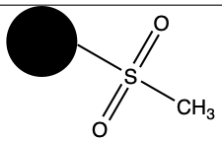
| Name | Structure | ΔMass (Da) |
|---|---|---|
| Cysteine sulphenic acid |  | +15.9949 |
| Cysteine sulphinic acid |  | +31.9899 |
| Cysteine sulphonic acid |  | +47.9848 |
| Methionine sulphoxide |  | +15.9949 |
| Methionine sulphone |  | +31.9899 |

**Table 3.1: Oxidation states of cysteine and methionine residues.** Adapted from Jeong *et al.* (2011).

### 3.8.1 Intact protein mass spectrometry

Based on the loss of post-transfer editing activity, evident from previous experiments (Section 3.6, AlaRS and ThrRS were incubated with $H_2O_2$ in a ratio of 1:10,000 and 1:200 (or water as a negative control) for 30 mins and 40 mins respectively. 50 µl samples containing 5-10 pmols purified protein in 50 % (v/v) methanol + 0.1 % (v/v) formic acid were analysed by electrospray mass spectrometry.

Figure 3.11a, 3.12a, 3.13a and 3.14a show molecular ions in the *m/z* range of 600-3000 for AlaRS (without $H_2O_2$), AlaRS (with $H_2O_2$), ThrRS (without $H_2O_2$) and ThrRS (with $H_2O_2$) respectively.

Following interpretation of the *m/z* spectra by the Maximum Entropy algorithm, a number of major species were identified around the expected molecular mass of AlaRS and ThrRS (97,566.12 Da and 75,829.58 Da respectively). For AlaRS, all

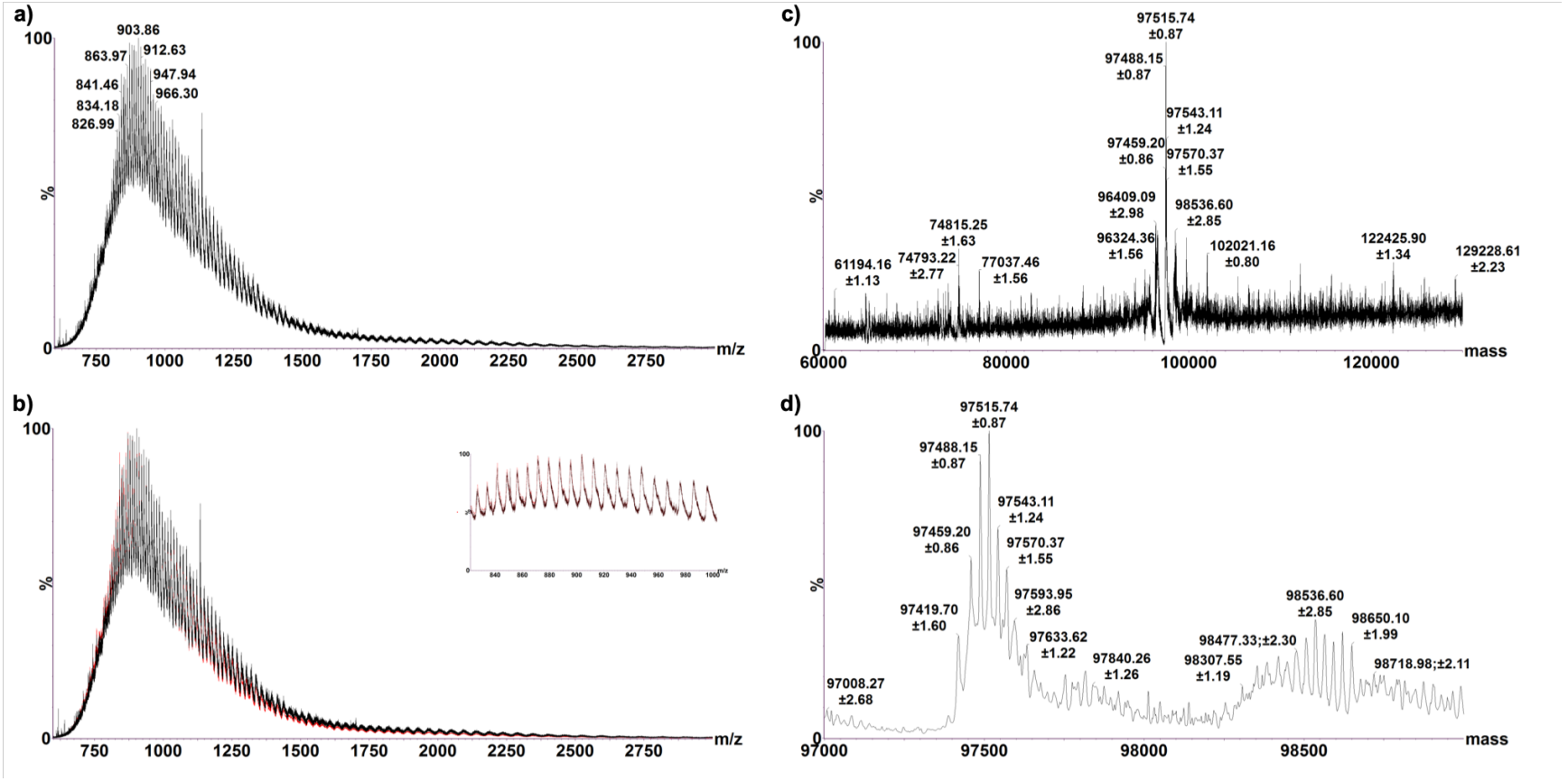



**Figure 3.11: Electrospray mass spectrometric analysis of AlaRS.** Electrospray mass spectrometry in positive ion mode. a) Electrospray mass spectrum b) Electrospray mass spectrum (black) overlaid with mock spectrum (red) c) Maximum Entropy interpretations of this spectra in the mass range 60,000 Da to 130,000 Da d) Deconvoluted Maximum Entropy interpretations of this spectra in the mass range 60,000 Da to 130,000 Da expanded between 97,000 and 99,000 Da.

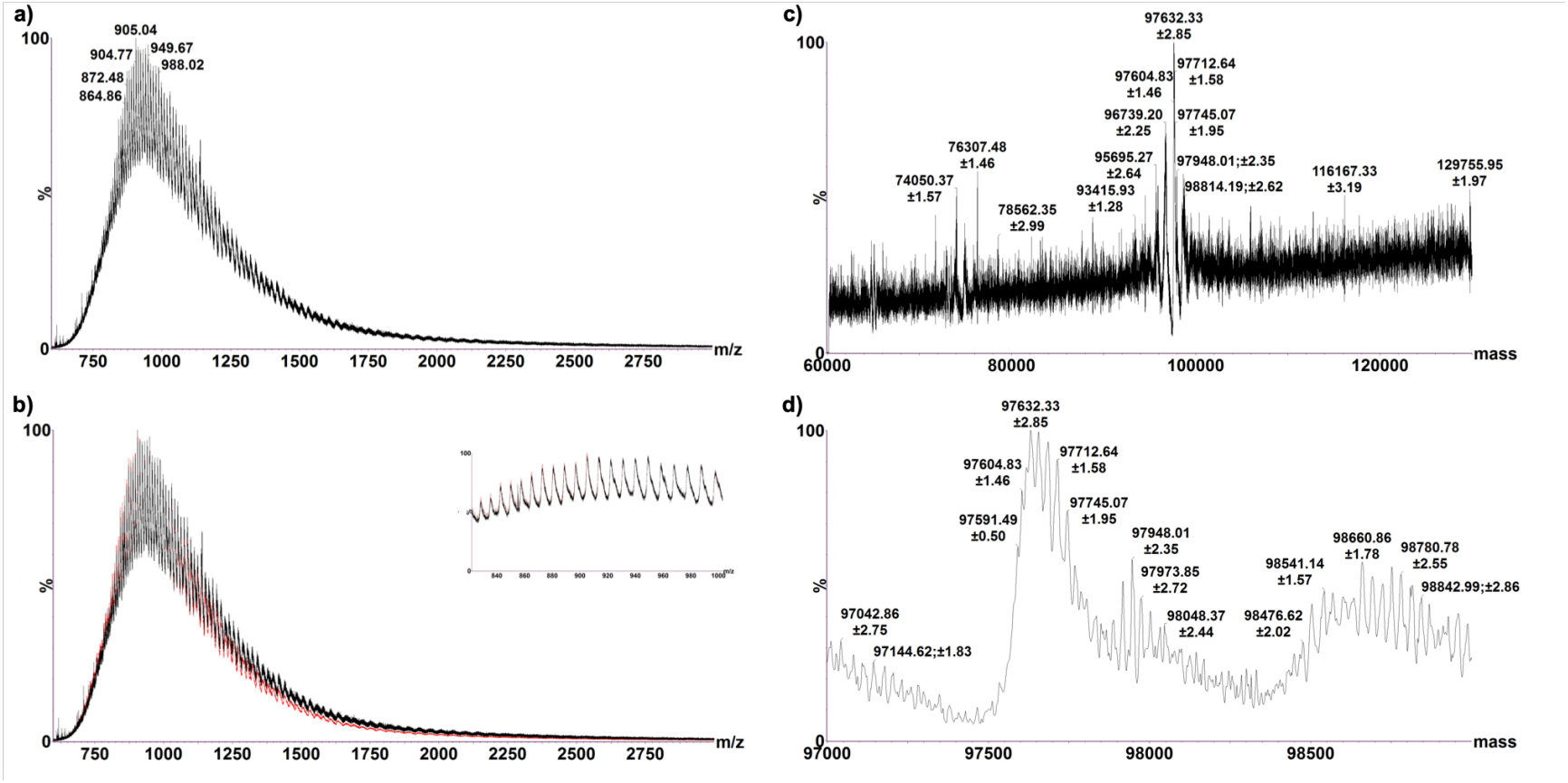



**Figure 3.12: Electrospray mass spectrometric analysis of AlaRS following incubation with $H_2O_2$ at a molar ratio of 1:10,000.** Electrospray mass spectrometry in positive ion mode. a) Electrospray mass spectrum b) Electrospray mass spectrum (black) overlaid with mock spectrum (red) c) Maximum Entropy interpretations of this spectra in the mass range 60,000 Da to 130,000 Da d) Deconvoluted Maximum Entropy interpretations of this spectra in the mass range 60,000 Da to 130,000 Da expanded between 97,000 and 99,000 Da.

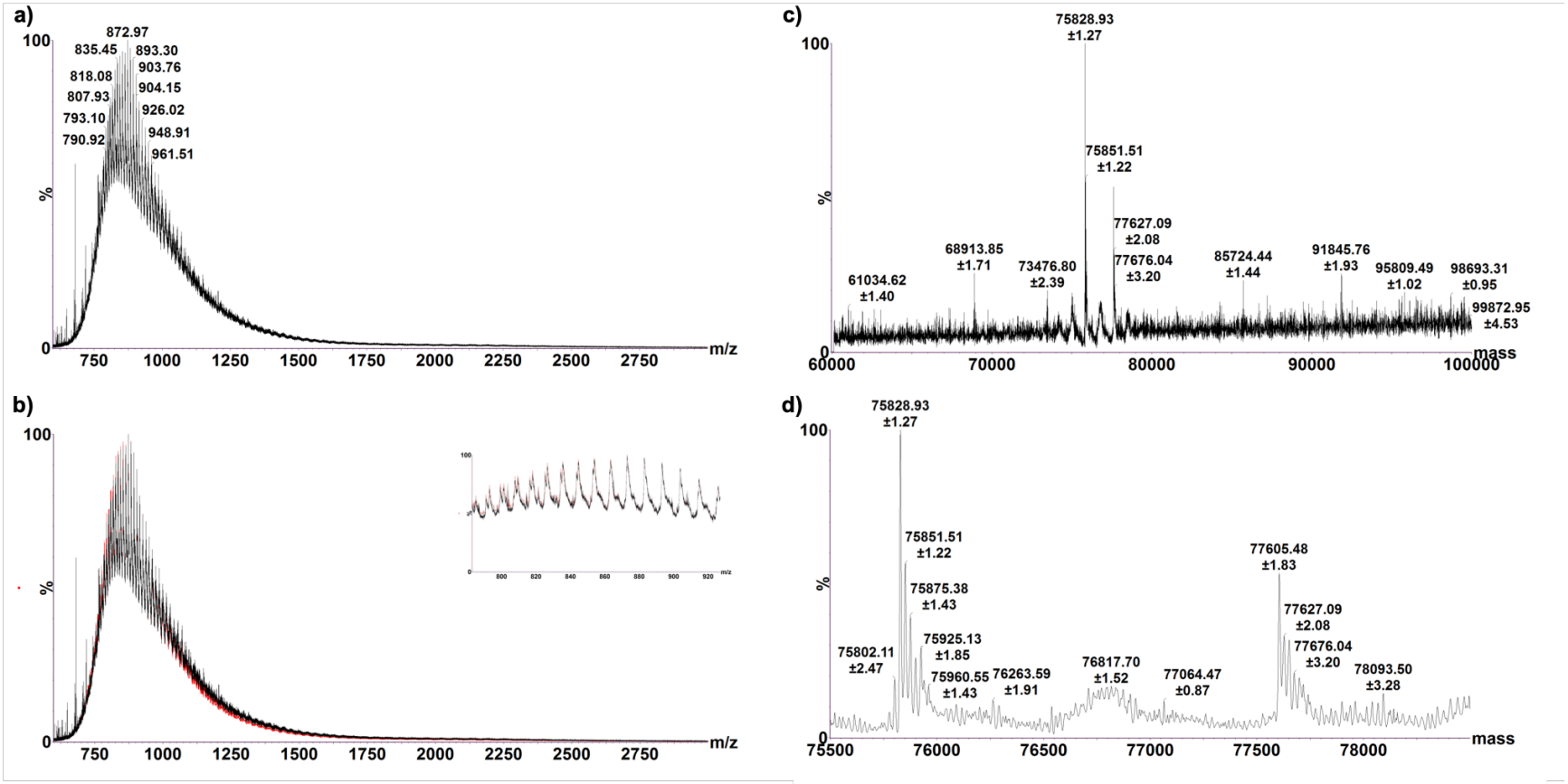



**Figure 3.13: Electrospray mass spectrometric analysis of ThrRS.** Electrospray mass spectrometry in positive ion mode. a) Electrospray mass spectrum b) Electrospray mass spectrum (black) overlaid with mock spectrum (red) c) Maximum Entropy interpretations of this spectra in the mass range 60,000 Da to 100,000 Da d) Deconvoluted Maximum Entropy interpretations of this spectra in the mass range 60,000 Da to 100,000 Da expanded between 97,500 and 78,500 Da.

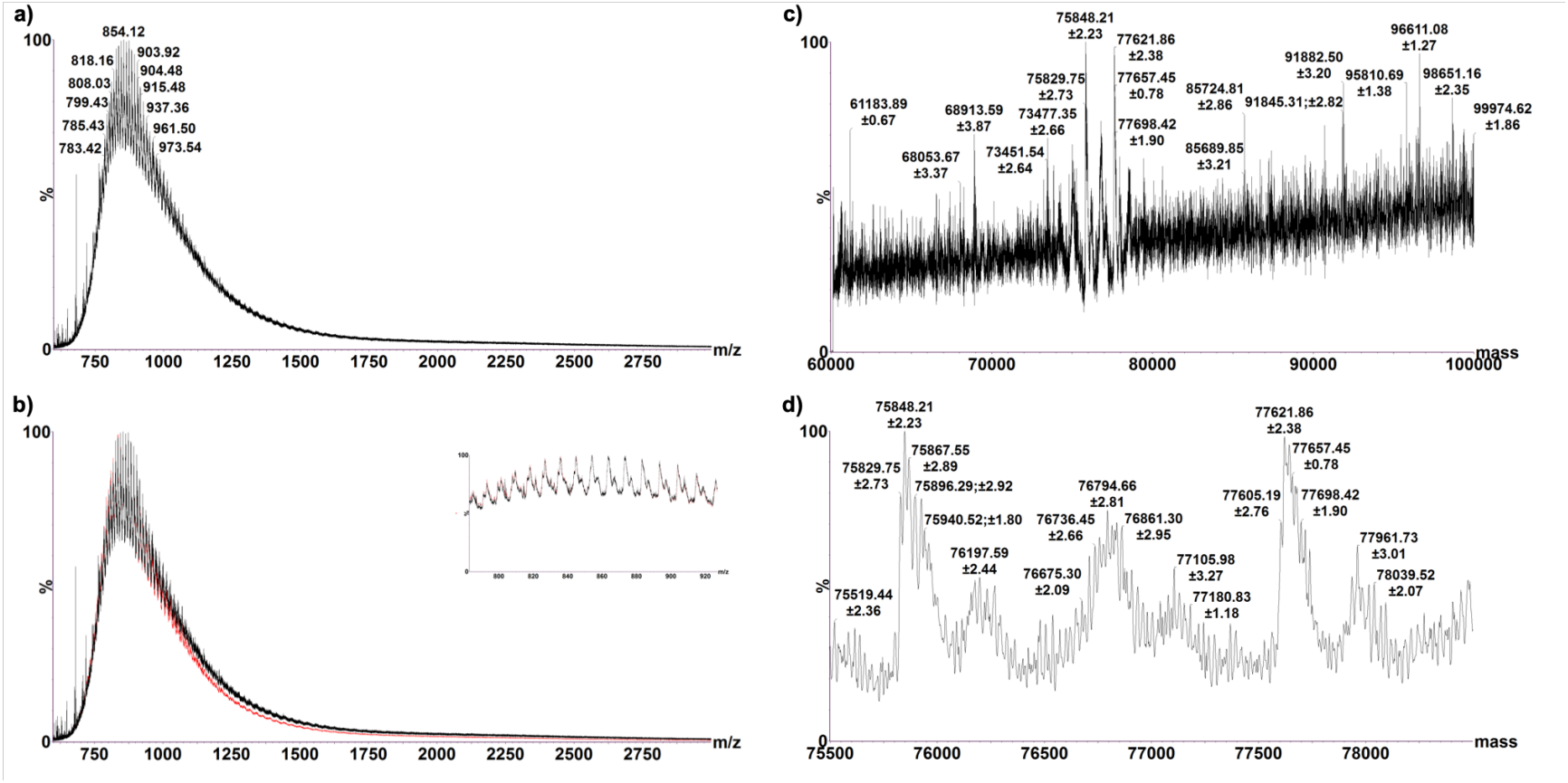



**Figure 3.14: Electrospray mass spectrometric analysis of ThrRS following incubation with $H_2O_2$ at a molar ratio of 1:200.** Electrospray mass spectrometry in positive ion mode. a) Electrospray mass spectrum b) Electrospray mass spectrum (black) overlaid with mock spectrum (red) c) Maximum Entropy interpretations of this spectra in the mass range 60,000 Da to 100,000 Da d) Deconvoluted Maximum Entropy interpretations of this spectra in the mass range 60,000 Da to 100,000 Da expanded between 97,500 and 78,500 Da.

species clustered at ∼97 kDa however in ThrRS two major clusters, ∼75 and ∼77 kDa were present. The presence of multiple clusters and species within each cluster may be explained by post-translational modifications (e.g. acylation, alkylation or phosphorylation) or loss of amino acids at the C- or N-terminus of the protein. In addition, in electrospray mass spectrometry, metal ion adducts are also common and so can generate modified species that are not found at the expected molecular mass. For example in Figure 3.11c, the peak with mass 97,459.20 Da may be attributed to AlaRS which has lost the N-terminal methionine and is sodiated (expected mass of 97,457.9 Da). Therefore, due to the complex nature of the protein samples this analysis did not aim to assign species to all major peaks but instead identify any shift between untreated AlaRS/ThrRS and the corresponding $H_2O_2$ treated samples a which indicates oxidation.

Figure 3.15 and 3.16 show that in both AlaRS and ThrRS respectively, mass spectra shifts to a higher molecular mass after treatment with $H_2O_2$, indicating that there is a stable modification occurring. The increase in mass is more apparent in AlaRS, an overall shift of ∼100 Da can be observed whereas in ThrRS the overall shift is less apparent ∼50 Da, this observation indicates that under these conditions AlaRS is more highly oxidised than ThrRS.

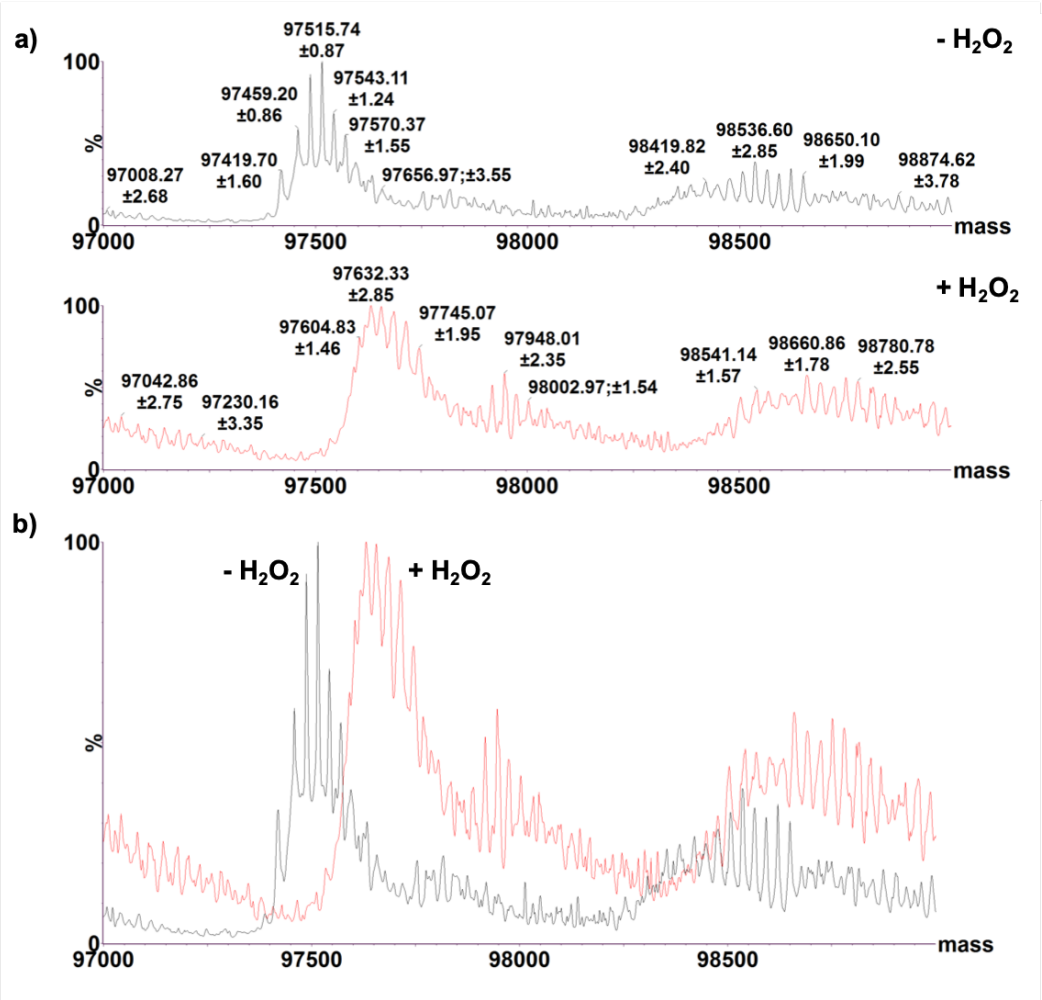



**Figure 3.15: Intact protein mass spectrometry of AlaRS with and without hydrogen peroxide treatment.**

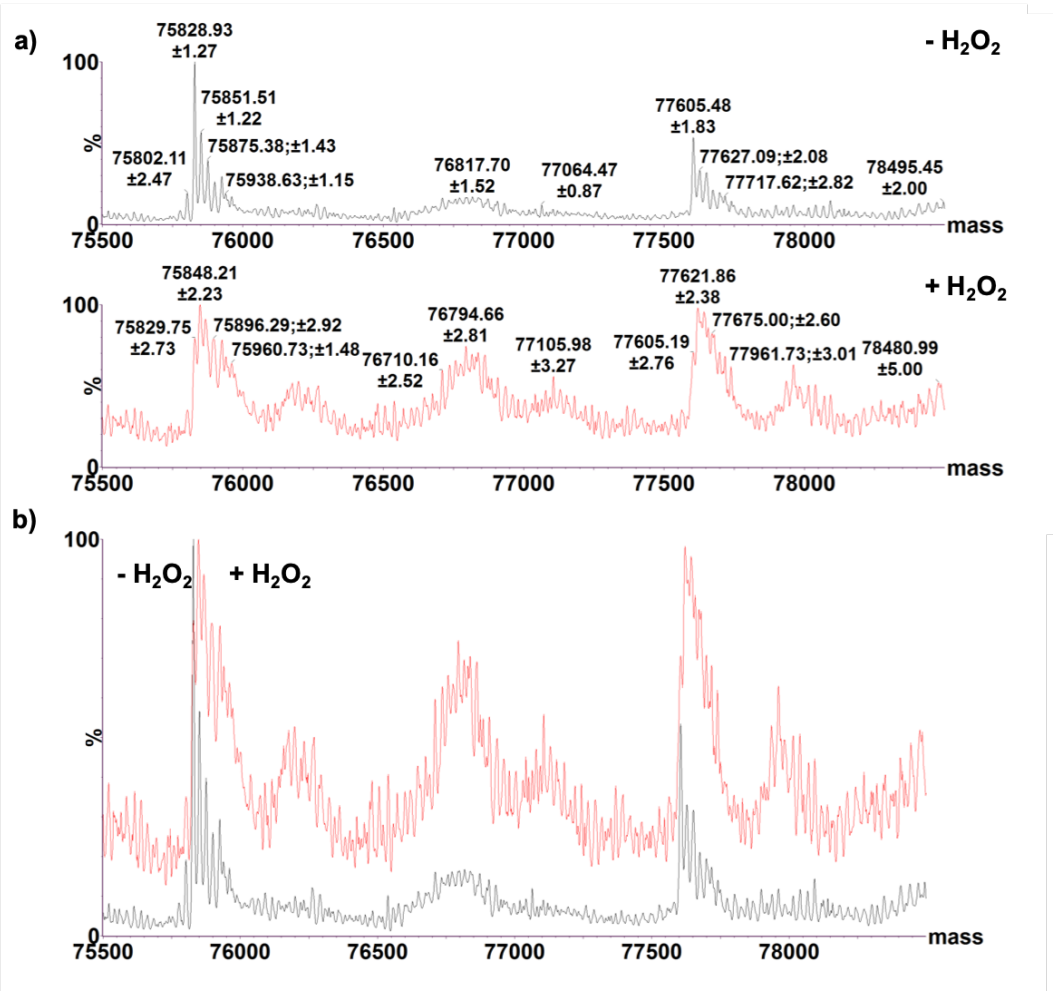



**Figure 3.16: Intact protein mass spectrometry of ThrRS with and without hydrogen peroxide treatment.**

### 3.8.2 Peptide-based mass spectrometry

AlaRS and ThrRS were incubated with $H_2O_2$ in a ratio of 1:10,000 and 1:200 (or water as a negative control), buffer exchanged to remove $H_2O_2$ and prepared for SDS-PAGE in a non-reducing loading buffer. Bands at a molecular weight corresponding to that of AlaRS and ThrRS were excised (Figure 3.17) and digested with trypsin for peptide-based mass spectrometry conducted by the WPH Proteomics RTP (University of Warwick).

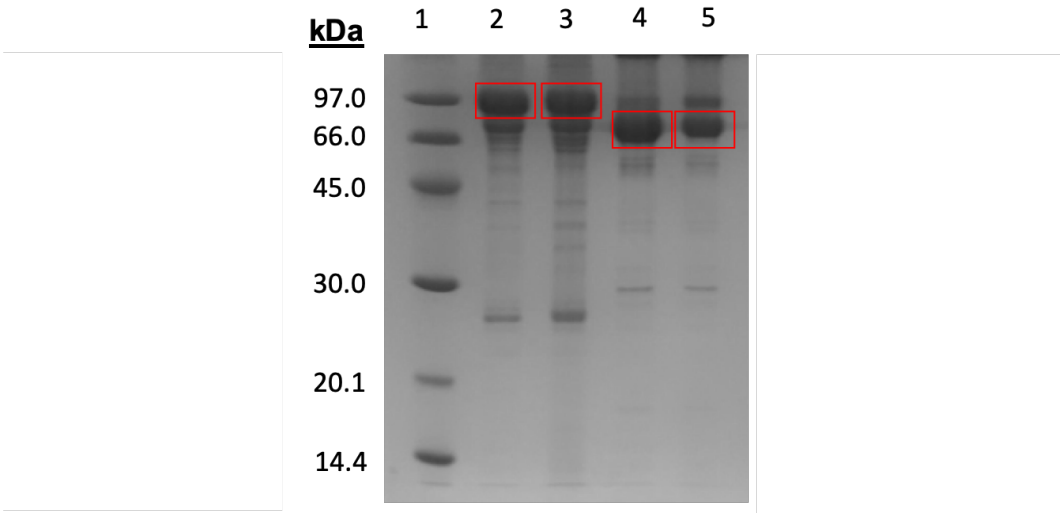



**Figure 3.17: 12 % SDS-PAGE of AlaRS and ThrRS samples for peptide based mass spectrometry.** Lane 1 - Amersham Protein Marker, Lane 2 - 15 µg AlaRS, Lane 3 - 15 µg AlaRS ($H_2O_2$ treated), Lane 4 - 15 µg ThrRS, Lane 5 - 15 µg ThrRS ($H_2O_2$ treated). Red boxes indicate bands which were excised for tryptic digest and mass spec analysis.

The mass of peptides corresponding to those containing modified residues, mono-oxidised (sulpenic acid), di-oxidised (sulphinic) and tri-oxidised (sulphonic acid) or unmodified residues, were identified.

For AlaRS, good sequence coverage was obtained; 74 % for the control sample and 77 % for the $H_2O_2$ treated sample. In each AlaRS sample up to 19 residues (methionine, tryptophan and cysteine) were found to be sites of potential modifications (Figure 3.18). Unexpectedly, Cys669 was tri-oxidised to sulphonic acid in both the control and treated AlaRS samples (Appendix 7.4). Sequence coverage was also good in ThrRS; 87 % for the control sample and 85 % for the $H_2O_2$ treated sample.

Each ThrRS sample contained 26 residues (methionine, tryptophan and cysteine) which contained modifications (Figure 3.19. Both Cys181 and Cys336 residues were tri-oxidised to sulphonic acid in the control and treated ThrRS samples (Appendix 7.4). In both AlaRS and ThrRS, all other modifiable residues were found as un-oxidised, mono-oxidised or di-oxidised forms. Most residues with the potential for modifications, were identified in multiple states, which indicates that the protein sample is heterogenous and many possible combinations of modifications exist within the sample. It is not possible to deconvolute between these different combinations to better understand the samples heterogeneity.

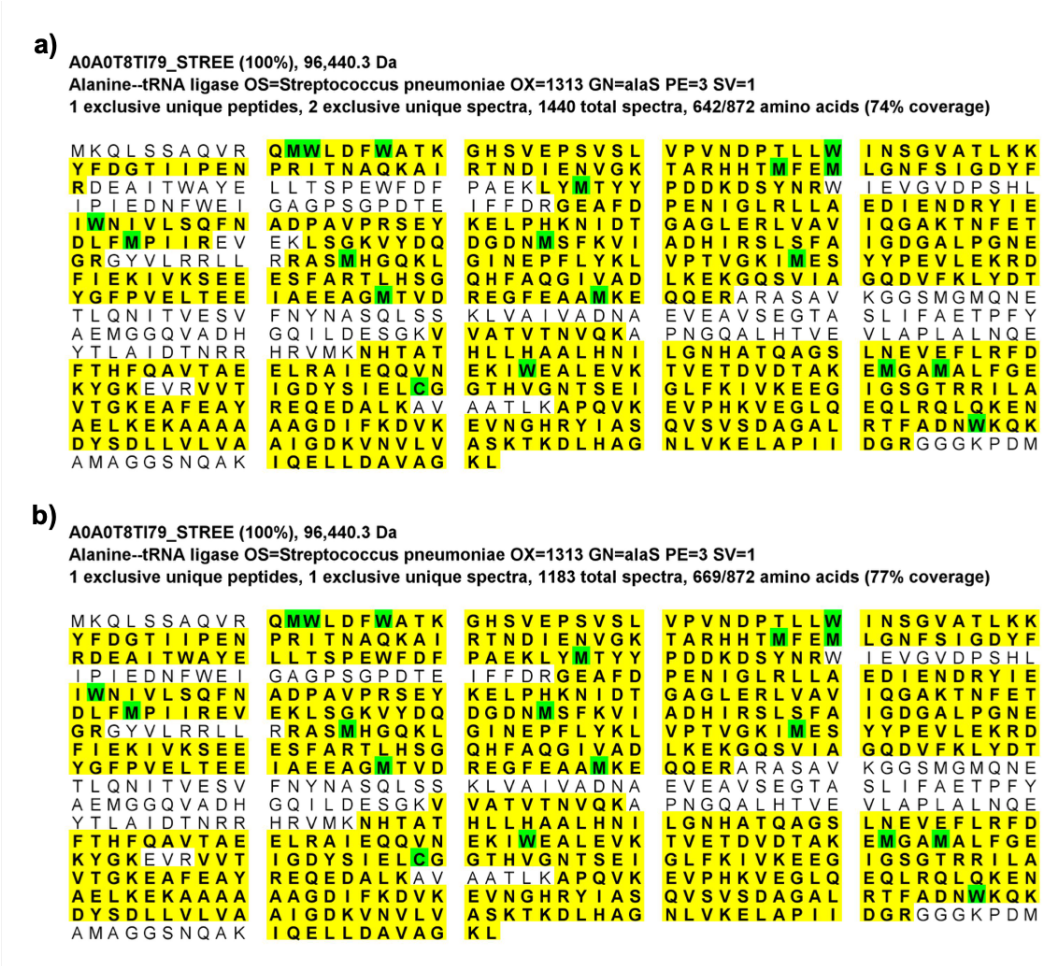



**Figure 3.18: Sequence coverage and modified residues of AlaRS.** a) AlaRS control sample, b) AlaRS $H_2O_2$ treated sample. Sequence coverage (yellow) and modified residues (green) are displayed. Results analysed in Scaffold (Version 4.8.1).

The high levels of oxidation and the presence of tri-oxidised cysteines in control

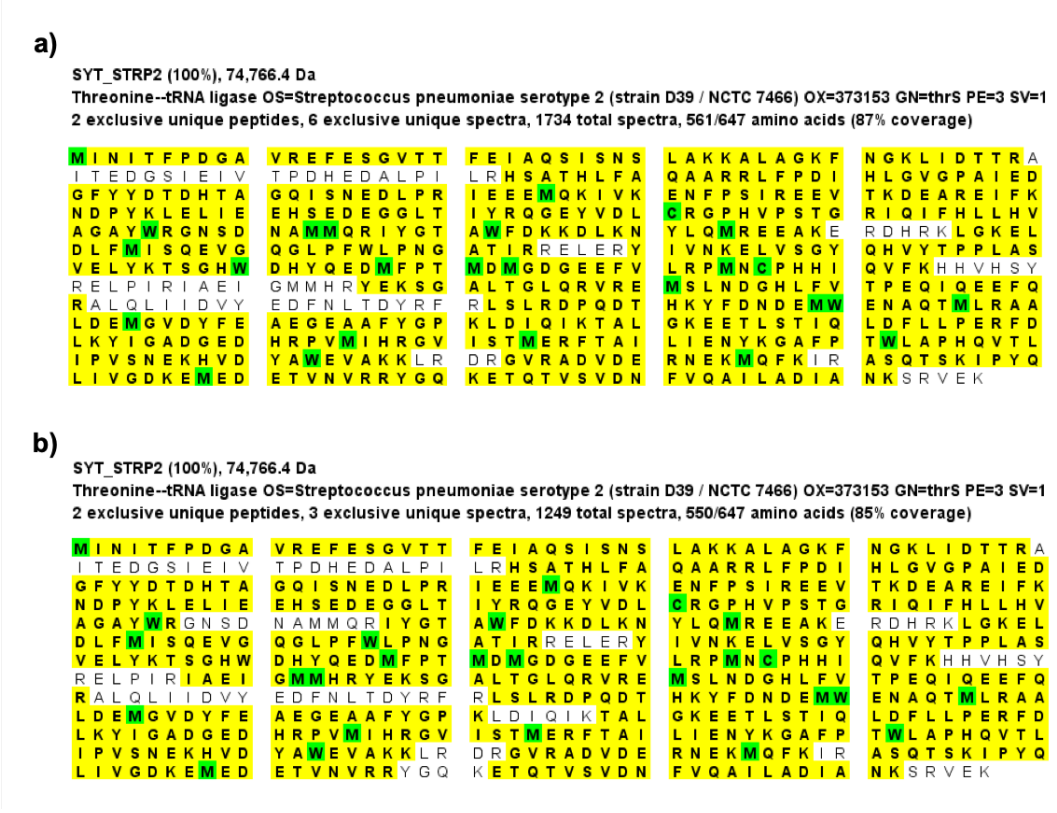



**Figure 3.19: Sequence coverage and modified residues of ThrRS.** a) ThrRS control sample, b) ThrRS $H_2O_2$ treated sample. Sequence coverage (yellow) and modified residues (green) are displayed. Results analysed in Scaffold (Version 4.8.1).

samples may have occurred as a result of sample preparation. Both SDS-PAGE and tryptic digest preparation occurs in an oxidative environment and it is common for methionine residues to show oxidation with peptide-based mass spectrometry. In order to prevent the reducing agent from reversing the oxidation from the $H_2O_2$ incubation, initial samples were prepared for SDS-PAGE in the absence of a reducing agent. Due to the large amount of oxidation observed, subsequent samples were prepared with reducing agent, however this did not appear to have any effect on the level or extent of oxidation in AlaRS or ThrRS.

## 3.9 Discussion

This chapter demonstrates that *S. pneumoniae* AlaRS and ThrRS are both capable of activating non-cognate amino acids leading to aminoacylation of tRNA with L-Ser. The presence of pre- and post-transfer editing mechanisms of AlaRS and ThrRS aim to minimise the impact of mischarging events, however, this research demonstrates that *in vitro* $H_2O_2$ can affect the ability of AlaRS and ThrRS to undertake post-transfer editing activity. $H_2O_2$ has been shown to have no effect on amino acid activation or pre-transfer editing activity but markedly reduces the rate of post-transfer editing in a concentration and time dependent manner. Intact protein mass spectrometry experiments have identified an increase in molecular mass of AlaRS and ThrRS species following incubation with $H_2O_2$ indicating that stable modifications to the proteins have occurred.

In addition, whilst both the amino acid activation and post-transfer editing activity of ThrRS was significantly lower than that of AlaRS, there was no increase in activity when the 6-His tag was transferred from the C-terminus to the N-terminus. Cleavage of the 6-His tag may improve ThrRS activity in future experiments.

Initially the background rate for the ADPNP amino acid activation assay (Figure 2.2) was intended to be obtained in the presence of the aaRS and absence of amino acid. However, it was observed that the initial rate of reaction was not consistent

or reproducible under seemingly identical conditions. The variation in initial rate was due to the variation in time between addition of the aaRS to the assay mix and the recording of spectrophotometric data. It was hypothesised that the aaRS's were therefore contaminated with trace amounts of IPP, a ubiquitous and highly active enzyme. To test this, the amino acid was added to the assay master mix and the first reaction was initiated by the addition of aaRS instead. Identical rates could be achieved if the reaction was initiated with either the aaRS or $PP_i$, but not amino acid, which confirmed that if aaRS was incubated with the $PP_i$, contaminating IPP would break down the $PP_i$ thereby affecting the initial velocity upon true initiation of the reaction.

Fluoride has previously been shown to inhibit IPP (Josse, 1966; Tono and Kornberg, 1967) and so 50 mM potassium fluoride (KF) was added to the assay mix, however this completely abolished ThrRS activity. It was therefore decided that for these experiments, the amino acid would be included in the assay master mix and the reaction would be initiated by the addition of the aaRS.

The amino acid activation assays were intended to test activity of the aaRS enzymes before progressing to further experiments. Whilst both AlaRS and ThrRS demonstrated activity in the ADPNP assay, the catalytic efficiencies of both proteins was low. The activity of ThrRS was $\sim$ 20 x lower than that of AlaRS and $>$5000 x less active than was reported for *E. coli* ThrRS (Ling and Söll, 2010; Ling *et al.*, 2012). These previous studies used a pyrophosphate exchange assay, which measures radioactivity transferred from pyrophosphate to ATP during amino acid activation. In contrast, experiments in this chapter utilised an alternate spectrophotometric assay to measure amino acid activation. Since the spectrophotometric assay uses an ATP analogue (ADPNP) it is possible that the greatly reduced kinetic constants were due to an inability of ThrRS to utilise this alternative substrate. The assay was repeated using another ATP homologue (ADPCP) which improved the observed amino acid activation rate only slightly. In addition to low amino acid activation activity, the concentration of ThrRS required to achieve an observable

and measurable rate of post-transfer editing was ∼20 x the concentration required for AlaRS to achieve a similar rate.

The requirement of zinc ions for ThrRS activity has been documented previously (Nureki *et al.*, 1993), but neither addition of zinc chloride ($ZnCl_2$) to the assay mixture or the purification protocol resulted in an improvement in enzymatic activity of ThrRS. It was identified that previous experiments with *E. coli* ThrRS were conducted with an N-terminal 6-His tagged protein (Ling and Söll, 2010), whilst in contrast the construct used in these experiments was a C-terminal 6-His tagged protein. This observation raised questions regarding the impact of the tag and importance of the C-terminus on ThrRS activity. However, ThrRS(N-term) showed no improvement in catalytic efficiency in either amino acid activation or post-transfer editing assays.

The effect of $H_2O_2$ on initial rates of pre- and post-transfer editing activities of AlaRS and ThrRS were demonstrated using a novel post-transfer editing assay developed by Dr Adrian Lloyd. In contrast to the assay used by (Ling and Söll, 2010), which was a radiolabelled end point assay measuring the production of mis-aminoacylated tRNA species, this assay is a continuous spectrophotometric assay, measuring the rate of pre- and post-transfer editing reactions by coupling the production of AMP during post-transfer editing to the oxidation of NADH to $NAD^+$ by myokinase, PK and LDH (Section 2.9.2).

The result from the post-transfer editing assays taken together show that oxidation of AlaRS and ThrRS is both a concentration and time dependent process *in vitro*. If AlaRS or ThrRS are in close proximity to significant amounts of $H_2O_2$ for extended periods of time in *S. pneumoniae*, oxidation and hence loss of post-transfer editing activity would likely occur *in vivo*. It is important to note that the concentrations of aaRS and $H_2O_2$ used in these assays were concentrations necessary to obtain accurately measurable rates of reaction and are not necessarily biologically relevant. The concentration of $H_2O_2$ and the effect of $H_2O_2$ on aaRS *in vivo* is likely to be more complex and the relative local concentrations of aaRS and $H_2O_2$ would be important.

A 35,000 x molar concentration of $H_2O_2$ is required to reduce the post-transfer editing rate of AlaRS by ∼50 %, whilst only 1,500 x molar concentration of $H_2O_2$ is required to reduce the post-transfer editing of ThrRS by ∼50 %. These values are in line with those previously reported by (Ling and Söll, 2010), whereby a significant reduction in total editing activity of 1.5 $\mu$M ThrRS was achieved with 4 mM $H_2O_2$ which is a 2,666 x molar concentration, and the lowest concentration of $H_2O_2$ reported to induce mis-aminoacylation of 2.3 $\mu$M ThrRS, was 200 $\mu$M or 86 x molar concentration.

$H_2O_2$ was found to have no effect on pre-transfer editing of AlaRS or ThrRS. As expected, the rate of pre-transfer editing of AlaRS was found to be higher for non-cognate amino acid when compared to the cognate amino acid. However, no significant difference in pre-transfer editing of ThrRS was observed between non-cognate amino acid and cognate amino acid. This may be explained due to the overall lower level of pre-transfer editing, and inherently larger error with ThrRS.

It is important to note that for all the post-transfer editing assays, since the reaction was initiated with amino acid and subsequently tRNA, aaRS (either with or without $H_2O_2$) was, in all cases added to the master mix prior to initiating and measuring the reaction. The time between the addition of aaRS to the master mix and the recording of data is referred to as zero, but is the minimum amount of time required for addition to and mixing with master mix, time for the spectrophotometer to be zeroed and time taken to obtain a suitable background rate. This duration is ∼40 seconds. Background rates were obtained for 1 minute prior to the addition of amino acid and 30 seconds following addition of amino acid prior to addition of tRNA. As such each recording has 2 - 2.5 mins of additional time for which aaRS and $H_2O_2$ were always incubated. This time was as consistent as possible and since data was obtained in triplicate any variation incurred from this has been accounted for in the measured error.

Intact protein mass spectrometry indicated that there was an increase in molecular mass of both AlaRS and ThrRS following treatment with $H_2O_2$. The quality of

this data was relatively poor and so increasing the protein concentration may have been beneficial. However, significant protein losses were incurred when a desalting column or centrifugal concentrator were used to remove the $H_2O_2$; as such protein concentration was the limiting factor. Additionally, as an alternative to direct injection, it may have been beneficial to use LC-MS (e.g. a C4 column). In *E. coli* ThrRS Cys182 and the equivalent Cysteine (Cys666) in AlaRS have been identified as essential for editing of misacylated amino acids (Ling and Söll, 2010; Beebe *et al.*, 2003). Since the editing domain of *S. pneumoniae* AlaRS contains the only cysteine residue (Cys669) in the protein, it was proposed that this residue is oxidised by $H_2O_2$ in a similar manner to that in *E. coli* ThrRS (Cys182) which resulted in loss of editing activity Ling and Söll (2010).

Peptide-based mass spectrometry was used in order to overcome the issues encountered with intact protein mass spectrometry. It was hoped that peptide-based mass spectrometry would identify oxidation of C669 and C181 of AlaRS and ThrRS respectively, and any additional residues which are oxidised upon exposure to $H_2O_2$.

However, analysis of protein modifications by mass spectrometry can be difficult, and a number of factors must be considered. Firstly, the mass shift in the peptide molecular weight, which in the case of oxidation is relatively small but can range from +16 Da to +48 Da and the stability of the modification are important. Secondly, the overall abundance of the modified peptide, and the effect of the modification on the peptides ionization efficiency are also factors to consider (Parker *et al.*, 2010).

Good coverage of both AlaRS and ThrRS was obtained, but many residues were highly oxidised in both the treated and untreated samples. This high level of oxidation may indicate that the loss of post-transfer editing activity is related to the proportion of protein oxidised at specific sites in the protein, rather than the absolute presence or absence of oxidation. Alternatively, the sample preparation, comprising SDS-PAGE, gel extraction and tryptic digest, which all occur in an oxidative environment, may have resulted in high levels of oxidation in these

proteins. Given that AlaRS and ThrRS samples are highly pure, the amount of oxidation occurring during sample preparation could be minimised by eliminating gel-extraction and preparing samples in solution for future mass spectrometric experiments.

Given that sequence coverage for both AlaRS and ThrRS is high, and that the residues of interest (cysteines) were successfully detected, this experiment acts as a proof of concept. With some optimisation, peptide-based mass spectrometry would be the preferred method for identification of specific sites which are modified and responsible for the loss of post-transfer editing activity of AlaRS and ThrRS. However, in order to thoroughly investigate these modifications a quantitative experimental design and more complex analysis is required. Since the intensity of the peak is unique to each peptide, comparison of the same peptide in its un-modified and modified form can be difficult and, in addition to $>$ triplicate runs the samples must be prepared and analysed at the same time, with equal loading of each sample. The ratio between modified and un-modified peptide can then be calculated for a specific sample, and this ratio can be compared to the equivalent ratio from another sample. This could provide information on whether the level of particular modifications varied between control and treated samples.

In addition to peptide-based mass spectrometry the relationship between loss of post-transfer editing and the oxidation of cysteine residues could be investigated using a 5,5-dithio-bis-(2-nitrobenzoic acid (DTNB) assay. Performing a DTNB assay after incubation of aaRS with $H_2O_2$ for different lengths of time, or at different $H_2O_2$ concentrations could determine if oxidation of cysteines occurred proportionally to the reduction in post-transfer editing activity.

## 3.10 Future Work

Specific modifications which are responsible for the loss of post-transfer editing activity observed in *in vitro* editing assays should be determined using quantitative

mass spectrometry experiments. Oxidation of Cys669 of AlaRS and Cys181 of ThrRS could also be investigated using a DTNB assay.

The work conducted here could also be extended to additional aaRSs. Ten aaRS families are capable of misactivating non-cognate amino acids, and therefore possess editing activities in order to maintain translational fidelity (Perona and Gruic-Sovulj, 2013). The effect of $H_2O_2$ on editing activities in aaRSs such as LysRS, which is known to misacylate $tRNA^{Lys}$ with both alanine and serine (Shepherd and Ibba, 2014), and ProRS which has been shown to misacylate $tRNA^{Pro}$ with alanine (Beuning and Musier-Forsyth, 2000) should also be investigated.

Additionally, the wider effects and consequences of $H_2O_2$ on misacylation of tRNA *in vivo* should be considered. The SpxB gene encodes a protein which, under aerobic conditions, converts pyruvate to acetyl phosphate, producing $H_2O_2$ as a metabolic bi-product (Spellerberg *et al.*, 1996). SpxB and MurM are both non-essential genes and so ΔSpxB mutants and ΔMurM mutants could be used to investigate the interplay between $H_2O_2$ and the PG *in vivo* (Regev-Yochay *et al.*, 2007). *S. pneumoniae* is a facultative anaerobe, therefore, the generation of $H_2O_2$ will depend on environmental $O_2$ concentrations. It is possible that aerobic growth creates the requirement for support by factors such as MurM to maintain the fidelity of protein synthesis; whilst anaerobic growth or growth without SpxB relaxes this necessity. If misacylated tRNAs, occurring as a response to increased $H_2O_2$ levels, are directed into PG synthesis by MurM, then it would be expected that the composition of the PG would alter in ΔSpxB or ΔMurM mutants. *S. pneumoniae (159)* shows a natural bias towards incorporation of alanine at the first position of the dipeptide bridge in PG. Therefore it would be expected that in *S. pneumoniae (159)* in the absence of SpxB and in the presence of MurM there would be higher proportions of alanine in the PG, and that in the presence of SpxB or exongenously added $H_2O_2$ the incorporation of serine into the the dipeptide bridge of PG would increase. Similarly, modulation of indirect cross-linking might also be apparent in wild type pneumococal peptidoglycan in cells grown aerobically compared to those grown

anaerobically. AlaRS or ThrRS editing site mutants, whereby the Cys669 and C181 respectively, have been mutated to serine would likely be lethal, but transformation with donor DNA containing this aaRS editing mutant may result in increased incorporation of serine into the dipeptide bridge. The composition of the PG layer can be investigated by mass spectrometry as described by Garcia-Bustos and Tomasz (1990).

*In vivo* the loss of post-transfer editing by synthetases in *S. pneumoniae* grown in aerobic conditions would result in a reduction in the fidelity of protein synthesis and ultimately cell death. Therefore, in the next chapter we investigate a proposed mechanism in which MurM acts to restore the fidelity of protein synthesis by preferentially incorporating amino acids from misacylated tRNA into the PG.

# Chapter 4

# Working towards characterising MurM kinetics and substrate specificity

The optimisation of iGln Lipid II synthesis was conducted in collaboration with Catherine Rowland (School of Life Sciences, University of Warwick). Next Generation Sequencing (NGS) of *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* was conducted at Micropathology Ltd with Daniel Hand providing experimental guidance and analysis, and John Moat (Antimicrobial Screening Facility, Warwick) for CAT2 training and MIC determination.

## 4.1 Introduction

MurM is an aminoacyl-tRNA-dependent ligase which is responsible for the addition of L-Ala or L-Ser to the third position lysine of the Lipid II(Lys) (Filipe *et al.*, 2000a). *MurM* belongs to the *MurMN* operon, where *MurN* encodes the protein responsible for the subsequent addition of a second amino acid, an invariable L-Ala, to that appended by MurM. MurM and MurN together generate a dipeptide

branched precursor required for the formation of indirect cross-links in the PG of *S. pneumoniae*. The *MurM* gene is required but not sufficient for the generation of high-level penicillin resistance. Deletion of the *MurM* gene has been shown to significantly reduce the $\beta$-lactam MIC of a strain, restoring susceptibility to resistant strains (Filipe *et al.*, 2001a). Therefore, MurM serves as a potential target for the development of novel antimicrobials, which could restore sensitivity to penicillin to otherwise resistant strains.

Whilst the proportion of cross-links in the PG varies very little between susceptible and resistant strains of *S. pneumoniae*, the level of indirect cross-linking between branched substrates, strongly correlates with resistance (Garcia-Bustos and Tomasz, 1990). This suggests that the ability of *S. pneumoniae* to generate indirect cross-links is closely related to the resistance of a strain. It was therefore proposed that remodelling of the PBP active site, which results in a reduced affinity for penicillin (and hence penicillin resistance) may also alter the substrate specificity, resulting in use of an alternate (branched) substrate (Garcia-Bustos and Tomasz, 1990). *MurM* is a highly mosaic gene, with many different MurM alleles. Whilst there is no particular MurM sequence required for resistance in *S. pneumoniae*, MurM alleles from a number of resistant strains contain a highly divergent sequence which differs in length and exact location, and is completely absent from susceptible strains (Filipe *et al.*, 2000b). This indicates that acquisition of this sequence is linked to the presence of resistance.

Studies by Lloyd *et al.* (2008) demonstrated that the catalytic activity of MurM was directly linked the observed proportion of branched muropeptide present in the pneumococcal cell wall of resistant and susceptible strains *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* respectively. The bias towards incorporation of L-Ser or L-Ala at the first position of the indirect dipeptide cross-link in the pneumococcal cell wall was not linked to the strains susceptibility, but was instead dictated by the specific MurM allele present. A 30 amino acid sequence from residue 244 to 274 was found to determine the specificity for L-Ala or L-Ser, with residue 260 being particularly

important. MurM alleles containing a lysine or threonine at position 260 were found to more commonly incorporate L-Ala or L-Ser respectively, into the pneumococcal cell wall (Filipe *et al.*, 2001b). Additionally, *in vitro* the amino acid selectivity of MurM was determined by the MurM allele and not due to strain specific differences in the tRNA pool (Lloyd *et al.*, 2008).

MurM alleles from *S. pneumoniae (Pn16)*, *S. pneumoniae (159)* and *S. pneumoniae (R6)*, which show differential bias towards amino acid incorporation to the PG *in vivo*, all showed a preference for mis-aminoacylated seryl-$tRNA^{Ala}$, compared to correctly-aminoacylated alanyl-$tRNA^{Ala}$ or seryl-$tRNA^{Ser}$ (Shepherd, 2011). This observation indicated that use of mis-aminoacylated tRNA substrates is of greater importance than the amino acid being incorporated. *S. pneumoniae* generates high levels of $H_2O_2$ as a biproduct of aerobic metabolism, and lacks many of the canonical ROS response mechanisms. As shown previously in Chapter 3, exposure of aaRSs to $H_2O_2$ results in a reduction in post-transfer editing and an increase in mis-aminoacylated tRNA. It was therefore proposed that MurM preferentially utilises amino acids from mis-aminoacylated tRNA species, encorporating them into the PG, in order to eliminate them from the cellular tRNA pool and thereby maintaining the fidelity of protein synthesis.

In addition, previous work on MurM demonstrating the selectivity of MurM for individual 'pure' substrates, was conducted using the $tRNA^{Ser3}$ isoacceptor. However, there are three isoacceptors for $tRNA^{Ser}$, which possess sequence differences at 51-63 bp which was shown previously to confer stability with EF-Tu (Sanderson and Uhlenbeck, 2007). We therefore questioned whether MurM utilises all isoacceptors equally, and aimed to characterise this before proceeding with experiments to understand MurM's preferences towards mis-aminoacylated tRNA's.

## 4.2 Aims

This chapter allowed preparation for a myriad of MurM experiments which will permit investigation of MurM substrate preference and kinetics. MurM from two strains (*S. pneumoniae (159)* and *S. pneumoniae (Pn16)*) was successfully sequenced, over-expressed and purified. Significant quantities of possible MurM substrates; Lipid II(Lys), iGln Lipid II(Lys), *S. pneumoniae (159)* tRNA and *in vitro* transcribed tRNA isoacceptors were (where necessary synthesised and) purified. Initial tests and protocol optimisation were conducted to ensure that enzymes and substrates were suitable for the proposed experiments. This preparative work will greatly assist future experiments which will investigate the ability of MurM to utilise mis-aminoacylated tRNAs and identify any difference in tRNA isoacceptor usage by MurM. In addition the Lipid II substrate for MurM will be determined, identifying where, with respect to MurM, amidation of the Lipid II precursor by MurT/GatD occurs.

## 4.3 Cloning, expression and purification of MurM from *S. pneumoniae (159)* and *S. pneumoniae (Pn16)*

### 4.3.1 Expression of $MurM_{159}$ and $MurM_{Pn16}$

Both $MurM_{159}$ and $MurM_{Pn16}$ have been cloned and expressed previously as per published methods (Lloyd *et al.*, 2008; Shepherd, 2011).

Constructs pET21b::$MurM_{Pn16}$(1) and pET21b::$MurM_{159}$(1) created by Lloyd *et al.* (2008) were used to transform *E. coli* C41(DE3)/pRIL cells. The proteins were expressed for 4 hours in accordance with the published protocol (Lloyd *et al.*, 2008). Despite no obvious signs of over-expression of MurM for either construct (Figure 4.1), the sample was progressed to purification due to reports of low levels of expression for MurM previously.

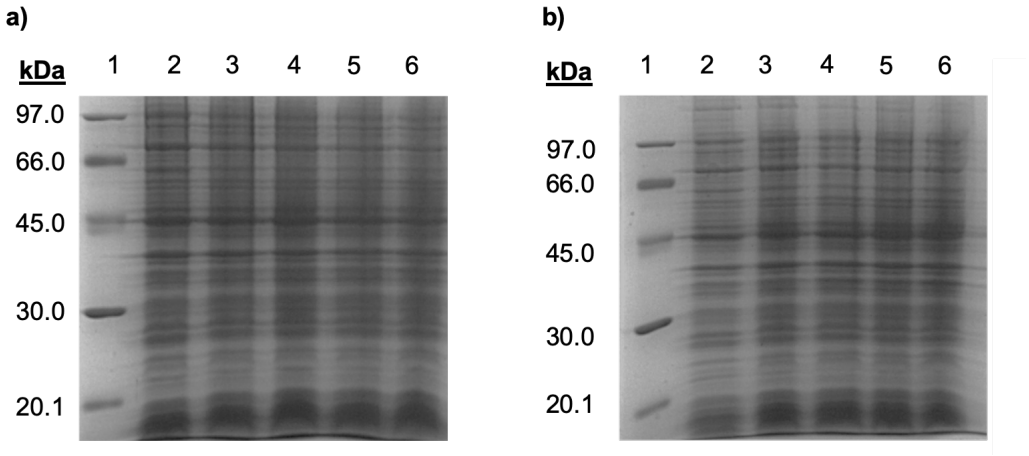



**Figure 4.1: 12 % SDS-PAGE showing 4 hour expression of MurM** a) $MurM_{Pn16}(1)$ and b) $MurM_{159}(1)$. Lane 1 - Amersham protein marker, Lane 2 - T=0, Lane 3 - T=1, Lane 4 - T=2, Lane 5 = T=3, Lane 6 - T=4. T = expression time (h).

In contrast to the numerous purification steps used by Lloyd *et al.* (2008), Shepherd (2011) had previously demonstrated that MurM could be purified in a one-step IMAC purification. Therefore purification of $MurM_{159}(1)$ using this method was attempted. Figure 4.2 shows that following IMAC, $MurM_{159}(1)$ contained a significant number of impurities and would require additional purification steps.

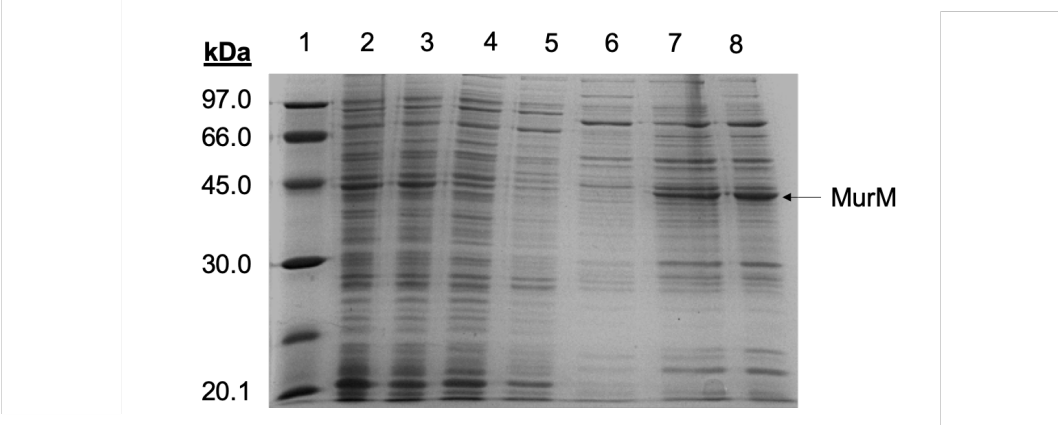



**Figure 4.2: 12 % SDS-PAGE showing IMAC purification of $MurM_{159}(1)$.** Lane 1 - Amersham protein ladder, Lane 2 - 10 µg protein from 10,000 *xg* supernatant, Lane 3 - 10 µg protein from column load, Lane 4 - 10 µg protein from column flow through, Lane 5 = 10 µg protein from column wash, Lane 6 to 9 - 20 µl of elution fractions.

The purification was therefore repeated in accordance with the protocol published by Lloyd *et al.* (2008). Figure 4.3 shows the fractions from the final IMAC purification step of this protocol. Using ProtParam the molecular weight of $MurM_{159}(1)$

(containing a 6-His tag) was predicted to be 48,500 kDa. A strong band at ~45 kDa, can be identified in the elution fractions, however, despite the improvement in purity compared to the one-step IMAC, there are still a large number of contaminating bands. The high levels of impurities were thought to be attributed to the low expression levels. Therefore optimising the expression level was deemed to be the most prudent way forward.

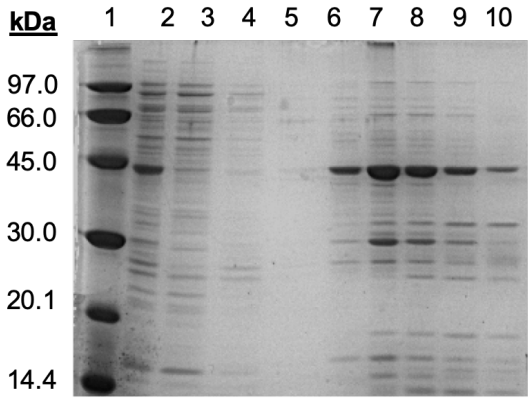



**Figure 4.3: 12 % SDS-PAGE showing final chromatography step (SEC) during purification of $MurM_{159}$(1).** Lane 1 - Amersham protein ladder, Lane 2 - 5 µg protein from column load, Lane 3 - 16 µl flow through, Lane 4 - 16 µl column wash, Lane 5 = 16 µl first elution fraction, Lanes 6 to 10 - 5 µg protein from elution fractions.

Constructs containing a TEV-cleavable 6-His tag (pET21b::$MurM_{159}$(2) and pET21b::$MurM_{Pn16}$(2)) were created by Shepherd (2011). Since the presence of a cleavable 6-His tag may prove useful in downstream experiments, and these constructs had been located, they were used from hereon in.

$MurM_{Pn16}$ was previously found to express slightly better than $MurM_{159}$ and so was the preferred choice for small scale expression trials in *E. coli* C41(DE3)/pRIL. The effect of temperature and IPTG concentration on MurM expression were investigated. Figures 4.4, 4.5 and 4.6 show both the insoluble (IS) and soluble (S) fraction of samples taken over 12 hours from expressions at 25 °C, 15 °C and 11 °C respectively. A band of ~45 kDa, which was not present at the 0 hour time point (T=0), can be seen at 2 hours (T=2) but it was difficult to determine whether or not a significant increase in expression was achieved over the time course.

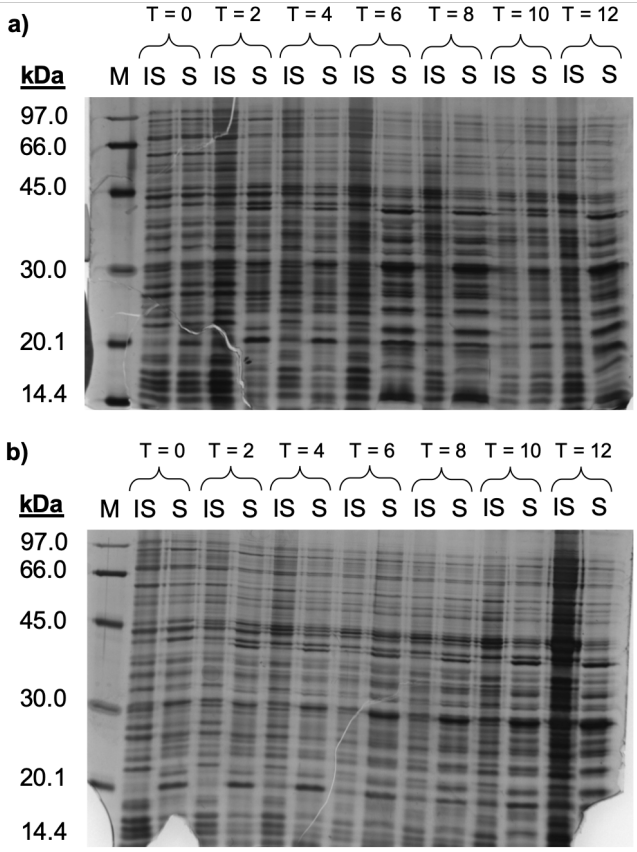



**Figure 4.4: 12 % SDS-PAGE showing a trial expression of $MurM_{Pn16}$(2) at 25 °C over 12 hours.** a) Expression induced by addition of 0.5 mM IPTG b) Expression induced by addition of 0.1 mM IPTG. 2 µg protein loaded per lane. T = time (hours), M = Marker (Amersham protein ladder), S = soluble fraction, IS = insoluble fraction. Gels were visualised by silver staining.

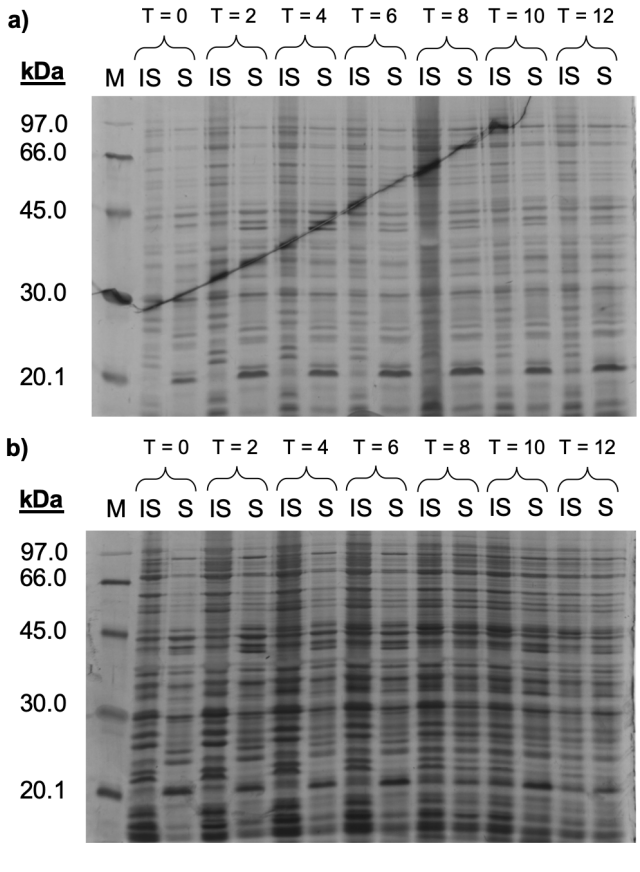



**Figure 4.5: 12 % SDS-PAGE showing a trial expression of $MurM_{Pn16}$(2) at 15 °C over 12 hours.** a) Expression induced by addition of 0.5 mM IPTG b) Expression induced by addition of 0.1 mM IPTG. 2 µg protein loaded per lane. T = time (hours), M = Marker (Amersham protein ladder), S = soluble fraction, IS = insoluble fraction. Gels were visualised by silver staining.

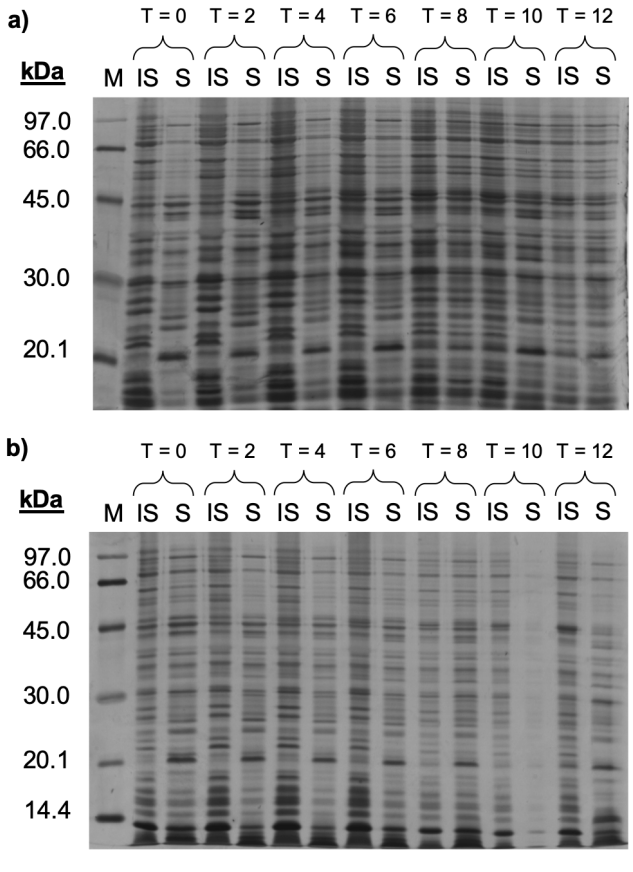



**Figure 4.6: 12 % SDS-PAGE showing a trial expression of $MurM_{Pn16}$(2) at 11 °C over 12 hours.** a) Expression induced by addition of 0.5 mM IPTG b) Expression induced by addition of 0.1 mM IPTG. 2 µg protein loaded per lane. T = time (hours), M = Marker (Amersham protein ladder), S = soluble fraction, IS = insoluble fraction. Gels were visualised by silver staining.

Conditions used in Figure 4.4 a) (25 °C and 0.5 mM IPTG) were used in a large scale expression for 4 hours. Samples at different time points were used to check expression levels by western blot (Figure 4.7).

The western blots showed the presence of the 6-His tag confirming that this protein was MurM. As previously reported, expression levels appeared to be higher in $MurM_{Pn16}(2)$ compared to $MurM_{159}(2)$. In addition, MurM is found in both the soluble and insoluble fractions, with a slightly higher proportion found in the insoluble sample.

Following an improved yield of MurM, $MurM_{159}(2)$ was purified in accordance with Lloyd *et al.* (2008) methods. Figure 4.8 shows SDS-PAGE analysis of fractions eluting from the final IMAC purification step. It is clear that whilst $MurM_{159}(2)$ is present at the expected molecular weight of ∼45 kDa a large number of contaminating proteins at both higher and molecular weights can be observed. Strong contaminating bands are present at a molecular weight of 30 kDa and 20 kDa.

Expression in LB generated a slight increase in yield of MurM and so was used in a subsequent large scale expression, cells were incubated at 37 °C for 2.5 hours and subsequently left for a further 12.5 hours at 25 °C. Cultures were disrupted using a cell disrupter and $MurM_{Pn16}(2)$ was purified by IMAC, reverse IMAC (following TEV cleavage) and SEC (Figure 4.9). Whilst the band corresponding to the molecular weight of $MurM_{Pn16}(2)$ was dominant, there were significant bands at both higher and lower molecular weights indicating contaminating proteins. A further hydrophobic interaction chromatography (HIC) purification step was applied to half the sample was shown to have no effect on the purity of this sample. The protein concentration was determined and the total yield was 19.3 mg, however, when tested in the mass spectrometry MurM activity assay this protein was found to be inactive.

Due to the ongoing issues with with low levels of expression, and the inability to obtain similarly pure or active protein according to previously published methods,

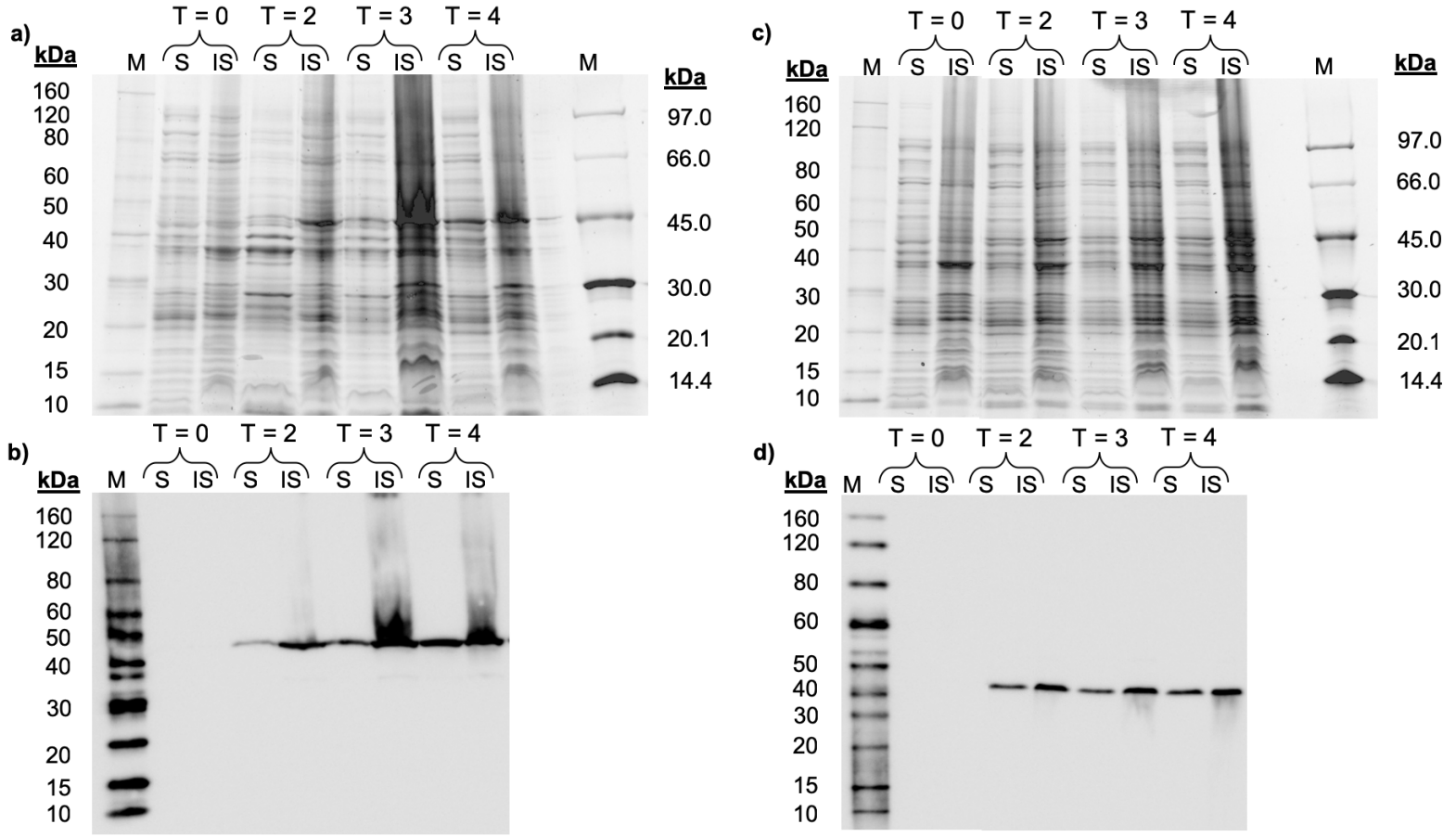



**Figure 4.7: 12 % SDS-PAGE and western blot analysis of MurM expression at 25 °C induced with 0.5 mM IPTG** a) SDS-PAGE of $MurM_{Pn16}$(2) b) Western blot of $MurM_{Pn16}$(2) c) SDS-PAGE of 9(2) d) Western blot of 9(2) 2 µg loaded per lane. T = time (hours), M = Marker (Bench mark His-tag standard and Amersham protein ladder), S = soluble fraction, IS = insoluble fraction.

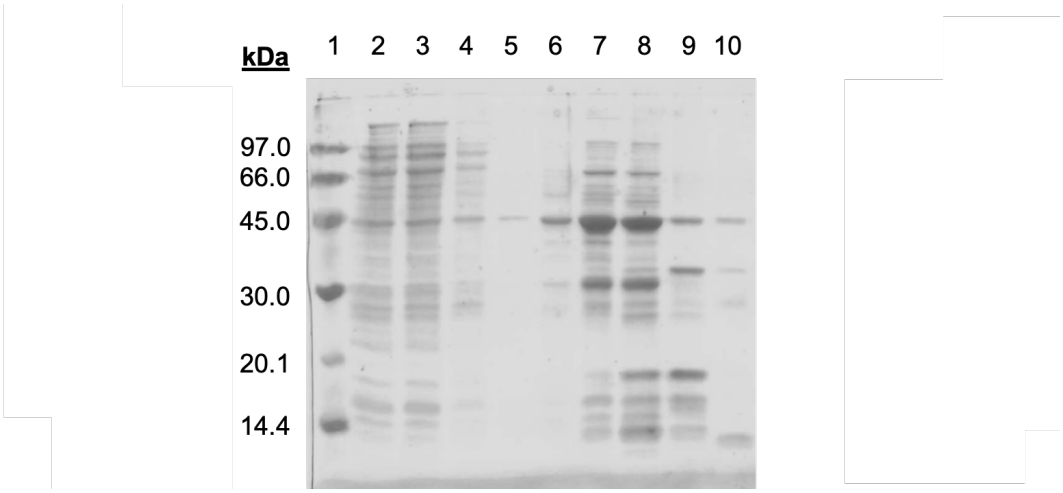



**Figure 4.8: 12 % SDS-PAGE showing final chromatography step (IMAC) during purification of $MurM_{159}$(2).** Lane 1 - Amersham protein marker, Lane 2 - 20 µl column load, Lane 3 - 20 µl flow through, Lane 4 - 20 µl column wash, Lane 5 to 10 = 20 µl of elution fractions.

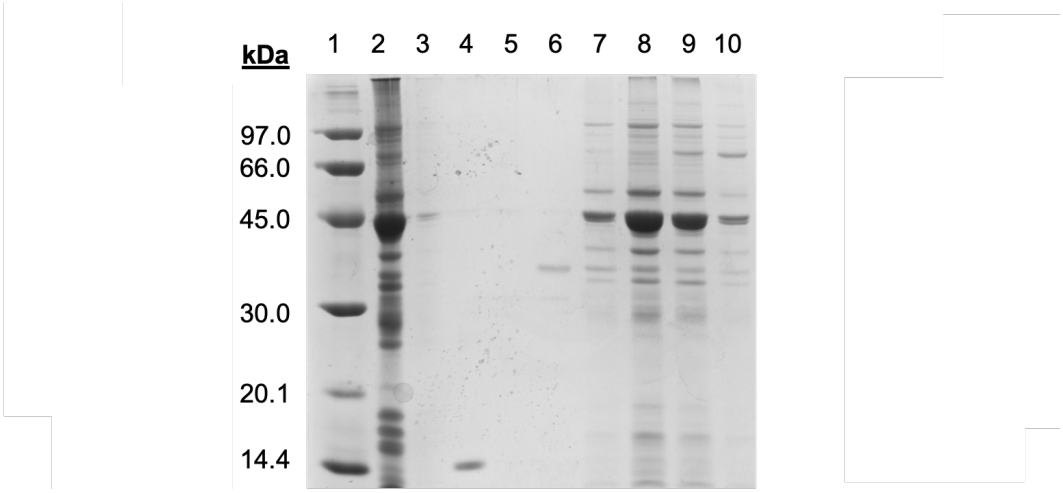



**Figure 4.9: 12 % SDS-PAGE showing final chromatography step (SEC) during purification of $MurM_{Pn16}$.** Lane 1: Amersham Marker, Lane 2 - 10 µl Sample Load, Lane 3 to 10 - 5 µl of elution fractions

it was deemed prudent to cease from using the constructs provided and to re-clone the MurM genes.

### 4.3.2 Cloning, expression and purification of MurM from *S. pneumoniae (159)* and *S. pneumoniae (Pn16)*

MurM from *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* were to be codon optimised and re-cloned, these new constructs would then undergo expression trials and purification; protocols would be based on previously published methods but optimised where necessary.

#### Gibson cloning of pET22b::MurM$_{Pn16}$ and pET22b::MurM$_{159}$

The sequence of *MurM*$_{Pn16}$ and *MurM*$_{159}$ (Lloyd *et al.*, 2008) were codon optimised and synthesised by IDT as g-blocks (referred to as *gMurM*$_{Pn16}$ and *gMurM*$_{159}$ from hereon in).

Successful amplification of *gMurM*$_{Pn16}$, *gMurM*$_{159}$ and pET22b are shown in Figure 4.10.

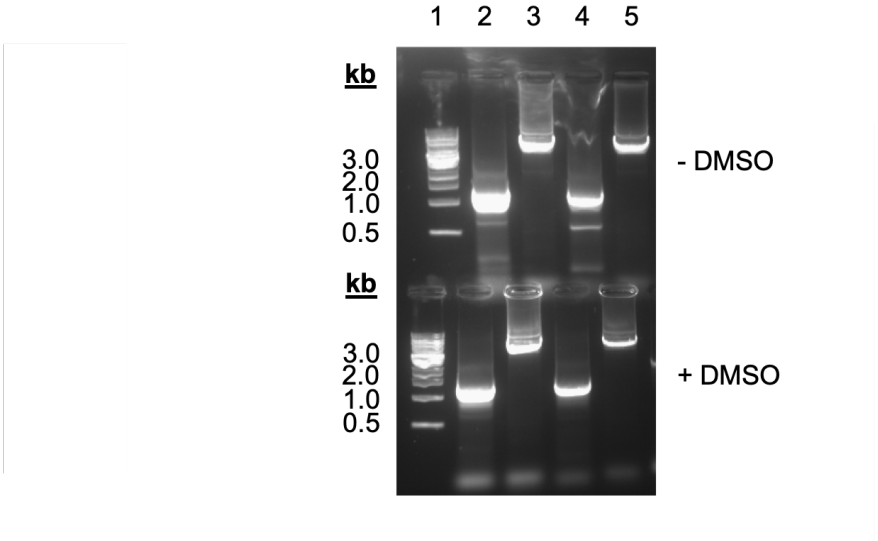



**Figure 4.10: 1 % agarose gel of *gMurM*$_{Pn16}$ and *gMurM*$_{159}$ and linearised pET22b** Lane 1: 1 kb DNA Ladder, Lane 2 - 50 µl *MurM*$_{Pn16}$, Lane 3 - 50 µl pET22b for *MurM*$_{Pn16}$, Lane 4 - 50 µl *gMurM*$_{159}$, Lane 5 - 50 µl pET22b for *gMurM*$_{Pn16}$

Following Gibson assembly and transformation into NEB-5$\alpha$ cells, 12 colonies were picked for each of pET22::gMurM$_{Pn16}$ and pET22b::gMurM$_{159}$. Following a colony PCR, 3 of the 12 colonies generated a band on a 1 % (w/v) agarose gel at $\sim$1 kb (corresponding to MurM). Plasmid was purified from these colonies and the presence of *gMurM$_{Pn16}$* or *gMurM$_{159}$* into the pET22b vector was confirmed by sequence analysis.

### 4.3.3 Expression of gMurM$_{159}$ and gMurM$_{Pn16}$

Small scale expression trials to compare expression of gMurM$_{Pn16}$ and gMurM$_{159}$ in *E. coli* BL21*R, C41(DE3) and C43(DE3) in both LB and AI media were conducted.

Figure 4.11 shows a strong band corresponding to the expected molecular weight of gMurM$_{159}$ in the insoluble fraction of BL21*R cells grown in LB media. There was little difference between 4 hour and overnight expression in LB, however expression in AIM was significantly lower. Smaller bands, indicating lower expression levels could also be observed in C41(DE3) cells but bands in C43(DE3) looked similar to those in the negative control.

Figure 4.12 shows a large band corresponding to the expected molecular weight of gMurM$_{Pn16}$ in the insoluble fraction of BL21*R cells grown in both LB and AIM. Significantly lower levels of overexpression of gMurM$_{Pn16}$ were achieved in C41(DE3) and C43(DE3) cells in both LB and AIM media.

Based on the trial expressions, the best conditions for over-expression of both gMurM$_{Pn16}$ and gMurM$_{159}$ were *E. coli* BL21*R, induced with IPTG for a 4 hour expression at 25 °C. gMurM$_{159}$ was transformed into BL21*R and expressed on a large scale (6 L LB) at the optimised conditions. gMurM$_{159}$ was purified as per (Lloyd *et al.*, 2008) and a strong band at $\sim$45 kDa, corresponding to the molecular weight of gMurM$_{159}$, can be observed throughout. A significant increase in purity was achieved between SEC and IMAC (Figure 4.13). The final purity of gMurM$_{159}$ was estimated to be $>$60 % using ImageJ (Version 1.51) and the total yield was 40

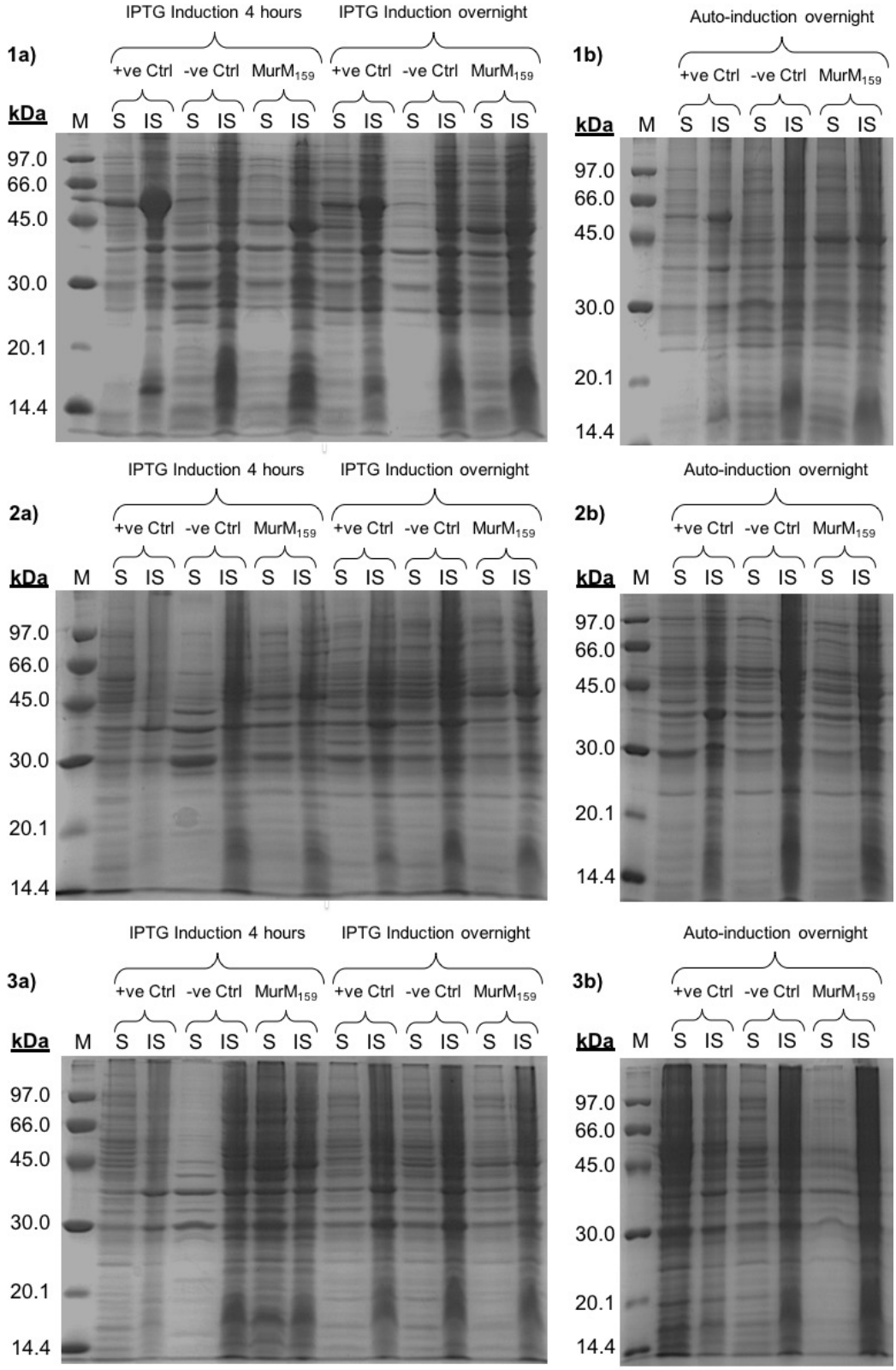



**Figure 4.11: 12 % SDS-PAGE analysis of expression trial of gMurM$_{159}$ in BL21*R, C41(DE3) and C43(DE3) in both LB media and AIM.** 1a) BL21*R in LB media, 1b) BL21*R in AIM, 2a) C41(DE3) in LB media, 2b) C41(DE3) in AIM, 3a) C43(DE3) in LB media, 3b) C43(DE3) in AIM. M = Amersham protein ladder, S = Soluble fraction, IS = Insoluble fraction

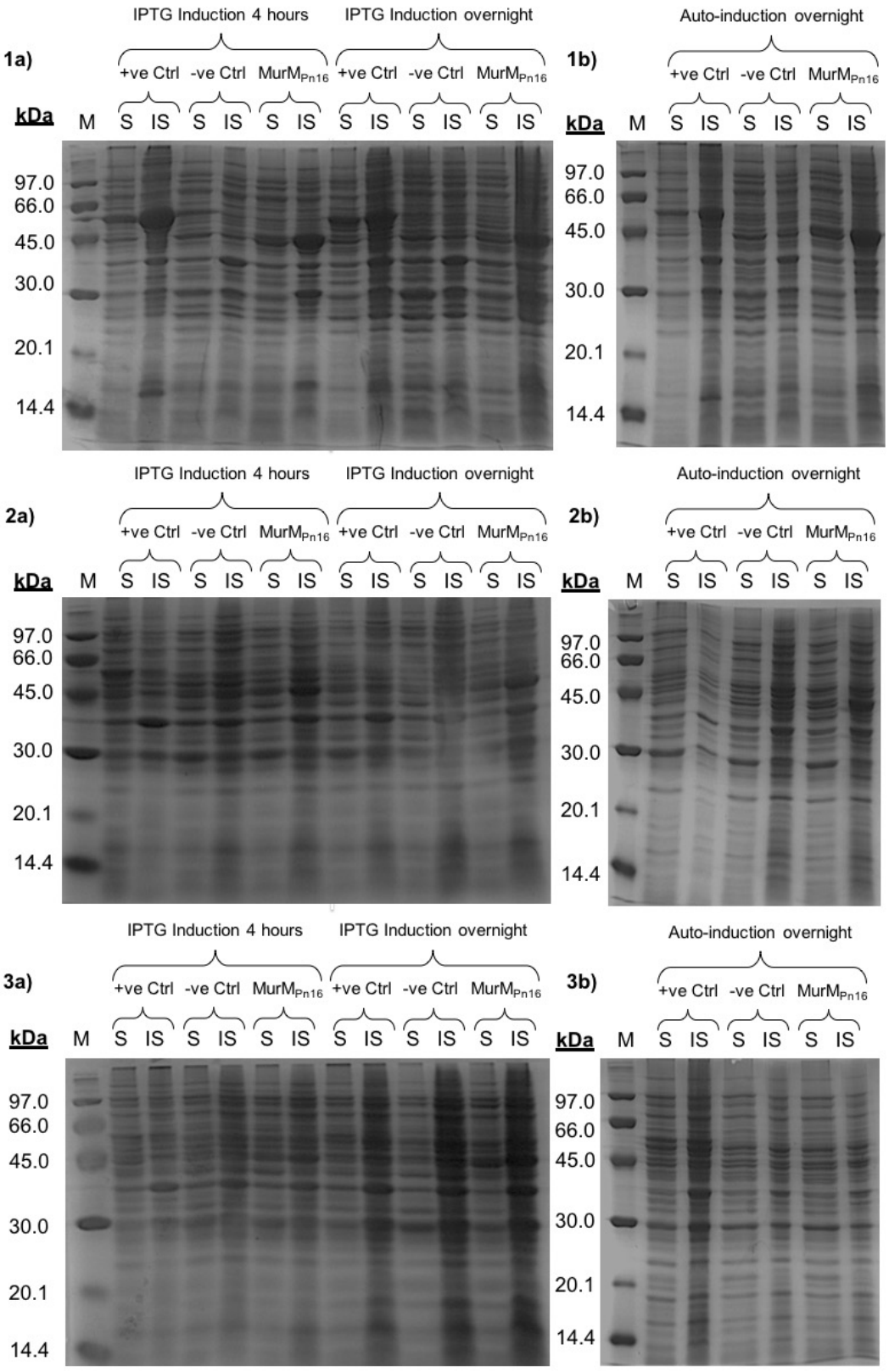



**Figure 4.12: 12 % SDS-PAGE analysis of expression trial of gMurM$_{Pn16}$ in BL21*R, C41(DE3) and C43(DE3) in both LB media and AIM.** 1a) BL21*R in LB media, 1b) BL21*R in AIM, 2a) C41(DE3) in LB media, 2b) C41(DE3) in AIM, 3a) C43(DE3) in LB media, 3b) C43(DE3) in AIM. M = Amersham protein ladder, S = Soluble fraction, IS = Insoluble fraction

mg.

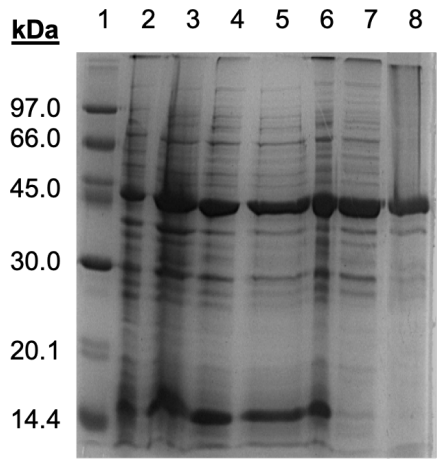



**Figure 4.13: 12 % SDS-PAGE showing gMurM$_{159}$ protein purification summary** Lane 1 - Amersham protein marker, Lane 2 - 20 µg protein from crude extract, Lane 3 - 20 µg protein from 10,000 *xg* pellet, Lane 4 - 20 µg protein from 100,000 *xg* supernatant, Lane 5 - 20 µg protein from ammonium sulphate cut at 25 %, Lane 6 - 20 µg protein from ammonium sulphate cut at 50 %, Lane 7 - 20 µg protein from SEC , Lane 8 - 20 µg protein from IMAC.

## 4.4 Characterising *S. pneumoniae (159)* and *S. pneumoniae (Pn16)*

A discrepancy of two amino acids, at position 134 and 135, between the gMurM$_{Pn16}$ sequence and the pET21(b)::MurM$_{Pn16}$ sequence was identified. The original pET21(b)::MurM$_{Pn16}$ clone, from which all previous work was conducted contained LQ at position 134 and 135, however the published sequence for MurM$_{Pn16}$ (Lloyd *et al.*, 2008) contained the amino acids MR at the corresponding positions. Since the gMurM$_{Pn16}$ and gMurM$_{159}$ had been synthesised based on the published sequences, it was deemed important to establish the correct sequence in order to ensure that the work conducted in this chapter is in line with previous experiments. In order to confirm that this was not a point mutation that had been accidentally introduced into pET21(b)::MurM$_{Pn16}$, the original *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* strains were located and sequenced. Locating the original strains was problematic and introduced new questions, as such it could not be confirmed that the strains that

were obtained were the exact *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* stocks required. Since the MIC's of these strains had previously been characterised and documented, and due to the characteristically high resistance of *S. pneumoniae (159)* to penicillin, the strains' MICs were first determined in an attempt to confirm their identity.

### 4.4.1 Susceptibility testing

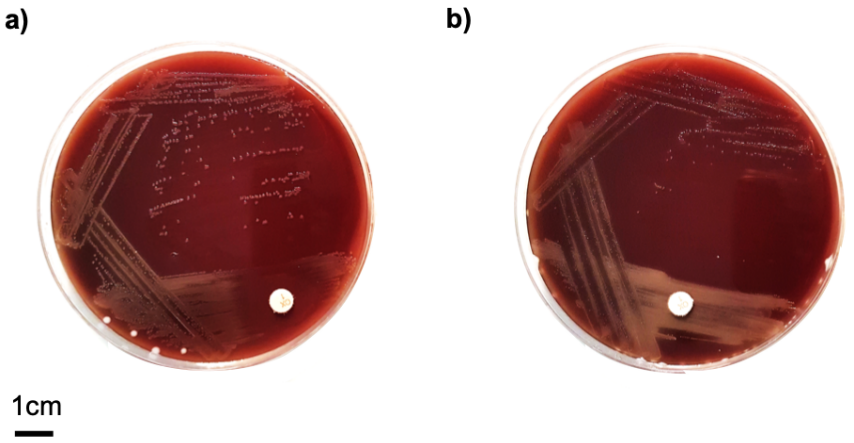



**Figure 4.14: Oxacillin Disk Susceptibility Testing.** Resistance of a) *S. pneumoniae (Pn16)* and b) *S. pneumoniae (159)* determined in line with the Clinical Laboratory Standards Institute (CLSI).

Initial attempts at classifying *S. pneumoniae (Pn16)* and *S. pneumoniae (159)* using oxacillin disk diffusion technique clearly highlighted the difference in susceptibility between the two strains (Figure 4.14). *S. pneumoniae (Pn16)* showed a zone of inhibition of ∼20 mm, whilst *S. pneumoniae (159)* did not show any zone of inhibition. Therefore according to EUCAST zone diameter breakpoints, *S. pneumoniae (Pn16)* and *S. pneumoniae (159)* are classified as susceptible and resistant strains respectively.

Figure 4.15 shows that MICs were determined to be 0.03 ng.µl$^{-1}$ and 32 ng.µl$^{-1}$ for *S. pneumoniae (Pn16)* and *S. pneumoniae (159)* respectively when using PenG E-strip method. The MIC of the positive control (ATCC *Staphylococcus aureus* 29213) was determined to be 0.25 ng.µl$^{-1}$ .

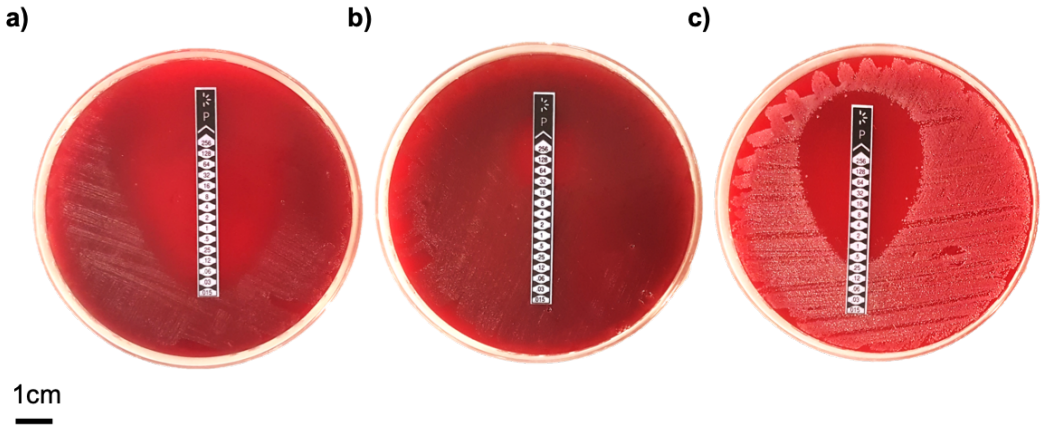



**Figure 4.15: PenG E-strip MIC Determination.** The MIC of a) *S. pneumoniae (Pn16)* and b) *S. pneumoniae (159)* was determined using PenG E-strips c) *ATCC Staphylococcus aureus 29213* was used as a positive control.

The E-strip and broth dilution techniques identified the MIC to be 32 ng.µl$^{-1}$ and 16 ng.µl$^{-1}$ respectively for *S. pneumoniae (159)* whilst for *S. pneumoniae (Pn16)* the MIC was determined to be 0.03 ng.µl$^{-1}$ for both methods. The MIC for the positive control (ATCC *Staphylococcus aureus* 29213) was determined to be 0.25 ng.µl$^{-1}$ and 0.2 ng.µl$^{-1}$ by E-strip and Broth Dilution respectively. The control MICs are in the middle of the EUCAST guidelines, as such the confidence in these results is high. The MICs obtained are similar to those determined previously using E-Test and agar dilution with benzylpenicillin (Barcus *et al.*, 1995). Despite the small variations in estimation of the MIC of *S. pneumoniae (159)* all MICs result in the classification of *S. pneumoniae (159)* as resistant and *S. pneumoniae (Pn16)* as sensitive according to EUCAST (EUCAST). The MICs obtained corroborated previous data and were therefore taken to be the *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* strains that were required.

### 4.4.2 Next generation sequencing

Due to the discrepancy between the published sequence and previous constructs sequences for  it was decided that both *S. pneumoniae (159)* and *S. pneumoniae (Pn16)*

strains MurM proteins should be re-sequenced, in order to establish the correct sequence for future studies.

Next generation sequencing (NGS) of the *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* genomes was deemed preferable to PCR of the MurM gene alone. NGS would generate a reliable MurM sequence as well as a full genome sequence for both strains, which may be benificial for making comparisons between these two strains. The NGS was conducted in collaboration with Catherine Rowland at Micropathology Ltd. with the guidance and supervision of Dr Daniel Hand.

De-novo assembly provided the full genome sequence of *S. pneumoniae (Pn16)* and *S. pneumoniae (159)*. MurM from *de novo* assembled *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* was aligned with the previously published MurM sequences (Figure 4.16). The NGS and *de novo* alignment confirms that the published sequence of $MurM_{159}$ is correct, however $MurM_{Pn16}$ is incorrect at amino acid position 134 and 135. The correct amino acid sequence at these positions is LQ, and is the protein sequence that has been used in all previous experiments. The MurM pET22b::$gMurM_{Pn16}$ therefore contains the two incorrect amino acids (MR) at these positions.

### 4.4.3 Site directed mutagenesis of $gMurM_{Pn16}$

The importance of the amino acids as position 134 and 135 on the MurM structure and activity is not known. Whilst the mutation at position 134 is conservative, at position 135 the mutation is non-conservative (a charged residue is mutated to a polar residue). Therefore it was decided that these positions of $gMurM_{Pn16}$ would be mutated from MR to LQ to keep experiments in line with all previous work.

GATC sequencing of $gMurM_{Pn16}$ in the pET22(b)::$gMurM_{Pn16}$ construct confirmed that the site directed mutagenesis had been successful. Position 134 and 135 in the amino acid sequence had been mutated from MR to LQ. From hereon in the mutated $gMurM_{Pn16}$ will be referred to as $gMurM_{Pn16}(1)$. Given that both $gMurM_{159}$

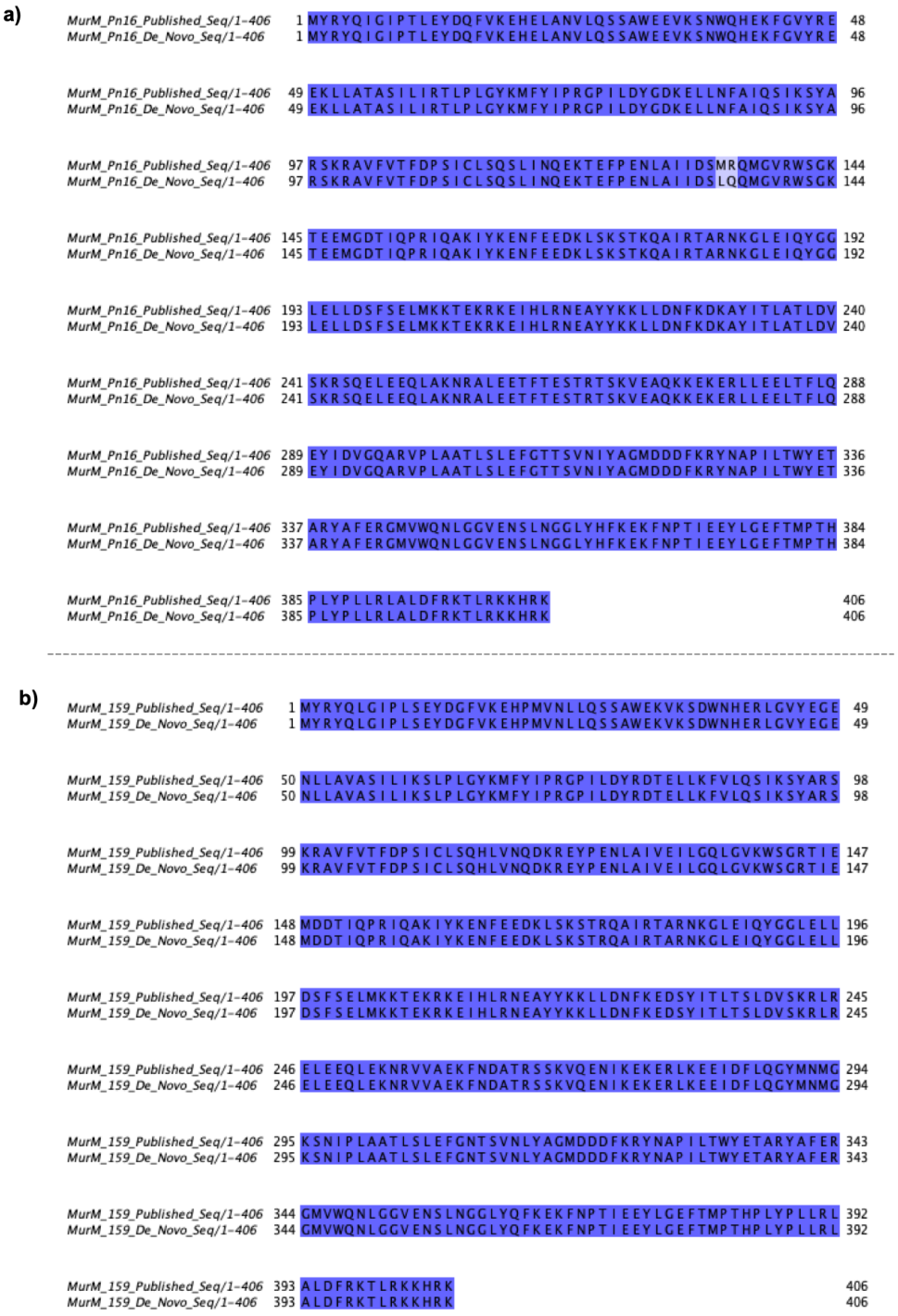



**Figure 4.16: Sequence alignments of MurM from *de novo* assembled a) *S. pneumoniae (Pn16)* and b) *S. pneumoniae (159)* with previously published MurM sequences.** Residues coloured according to % identity using Jalview (Version 2.10.5).

(containing LQ at position 134 and 135) and gMurM$_{Pn16}$ (containing MR at position 134 and 135) both showed optimal over-expression in the same conditions, the trial expression was not repeated with gMurM$_{Pn16}$(1) and the previously established expression conditions were used.

## 4.5 Lipid II(Lys) and iGln Lipid II(Lys) synthesis

In order to identify the true substrate of MurM both Lipid II(Lys) and iGln Lipid II(Lys) were synthesised. Synthesis and purification procedures for all Lipid II(Lys) or iGln Lipid II(Lys) substrates are outlined in Section 2.5-2.7. Structures and details such as molecular weights of intermediate and final compounds are shown in Appendix 7.5.

### 4.5.1 Synthesis of UDP-Mur*N*Ac-5P and iGln UDP-Mur*N*Ac-5P

**Expression and Purification of Mur ligases A-F**

Mur ligases A-F were expressed and purified by IMAC, the final purity of the Mur ligases was >75 % Figure 4.17.

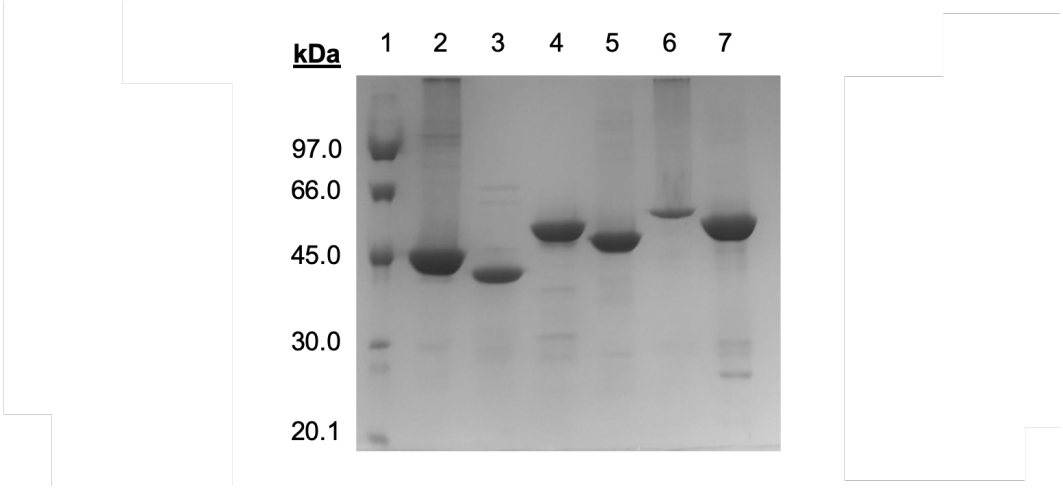



**Figure 4.17: 12 % SDS-PAGE showing final purity of Mur ligases A-F.** Lane 1 - Amersham protein marker, Lane 2 - 20 µg MurA, Lane 3 - 20 µg MurB, Lane 4 - 20 µgMurC, Lane 5 - 20 µg MurD, Lane 6 - 20 µg MurE, Lane 7 - 20 µg MurF.

**UDP-Mur*N*Ac-5P and iGln UDP-Mur*N*Ac-5P synthesis**

Both UDP-Mur*N*Ac-5P and iGln UDP-Mur*N*Ac-5P were synthesised and purified to a purity of >97 % (Figure 4.18). UDP-Mur*N*Ac-5P and iGln UDP-Mur*N*Ac-5P eluted at a conductivity of 25.62 mS/cm and 13.51 mS/cm as expected. The successful synthesis of UDP-Mur*N*Ac-5P and iGln UDP-Mur*N*Ac-5P were confirmed by negative ion mass spectrometry (Figures 4.19 and 4.20).

### 4.5.2 Optimisation of iGln Lipid II(Lys) synthesis

Due to previously low yields of iGln Lipid II(Lys), optimisation of the synthesis and purification of iGln Lipid II(Lys) was pursued.

Lipid II(Lys) synthesis was used as a starting point for iGln Lipid II(Lys) synthesis. The effects of doubling the iGln UDP-Mur*N*Ac-5P concentration, incubation overnight (compared to the standard 3 hour incubation) and addition of moenomycin were investigated. Figure 4.21 indicates that doubling UDP-Mur*N*Ac-5P concentration to 4 mM and incubation at 37 °C overnight, increased iGln Lipid II(Lys) yields. However, incubation with moenomycin had no effect, suggesting that there was no significant loss of iGln Lipid II(Lys) product arising from TG activity of *M. flavus* PBPs. It was noted that Lanes 5 and 8 ran differently to all other lanes, however the cause of this remains uncertain (Figure 4.21).

A time course of incubation over 13.5 hours was conducted to ensure that iGln Lipid II(Lys) was harvested at the optimal time. Figure 4.22 shows that whilst there is a steady increase in yield over the first 6 hours, there is no further increase after this time point. However, since there was no decrease in product when the syntheses were incubated for up to 13.5 hours, overnight syntheses were used for both Lipid II(Lys) and iGln Lipid II(Lys) for convenience.

The iGln Lipid II(Lys) purification was based upon that of Lipid II(Lys), whereby

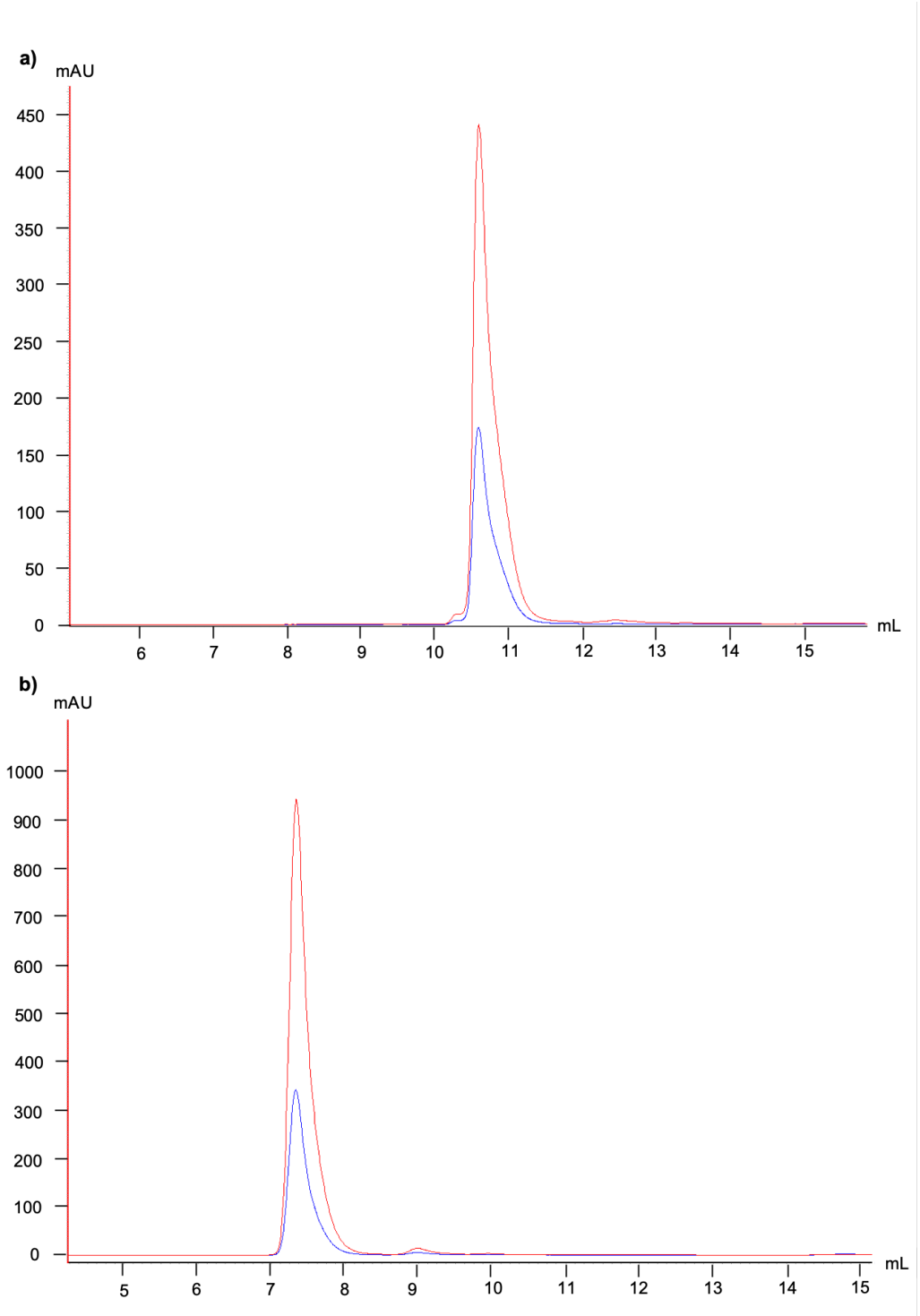



**Figure 4.18: Chromatograms from MonoQ Purity Test of a) UDP-Mur*N*Ac-5P and b) iGln UDP-Mur*N*Ac-5P.** Absorbance at 254 nm (Red) and absorbance at 280 nm (Blue) recorded. UDP-Mur*N*Ac-5P and iGln UDP-Mur*N*Ac-5P eluted at a conductivity of 25.62 mS/cm and 13.51 mS/cm respectively.

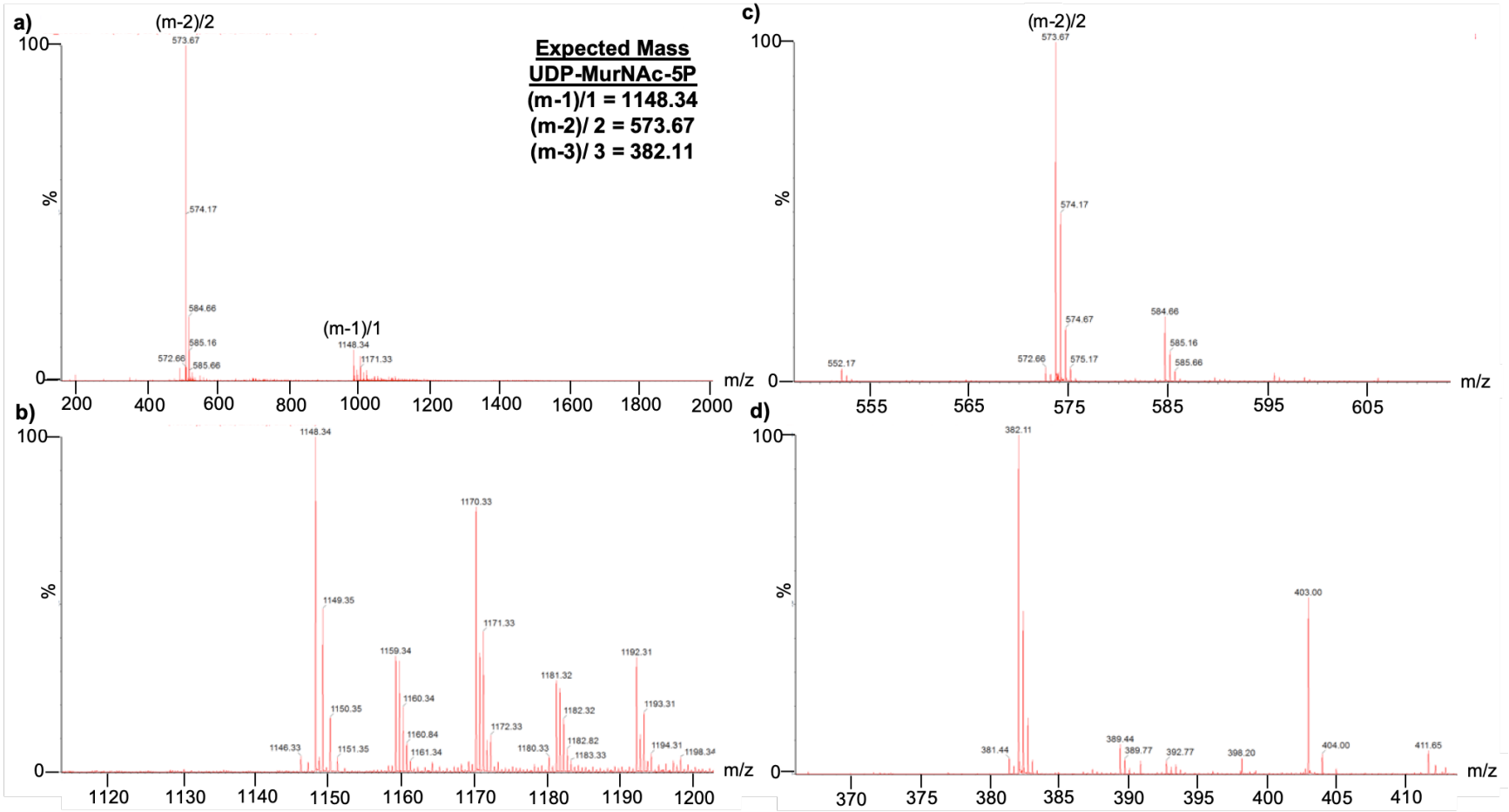



**Figure 4.19: Negative ion ES-MS of UDP-MurNAc-5P** a) Full spectra, b) Enlarged spectra showing the series corresponding to (m-1)/1 (1148.34) c) Enlarged spectra showing the series corresponding to (m-2)/2 (573.67) d) Enlarged spectra showing the series corresponding to (m-3)/3 (382.11).

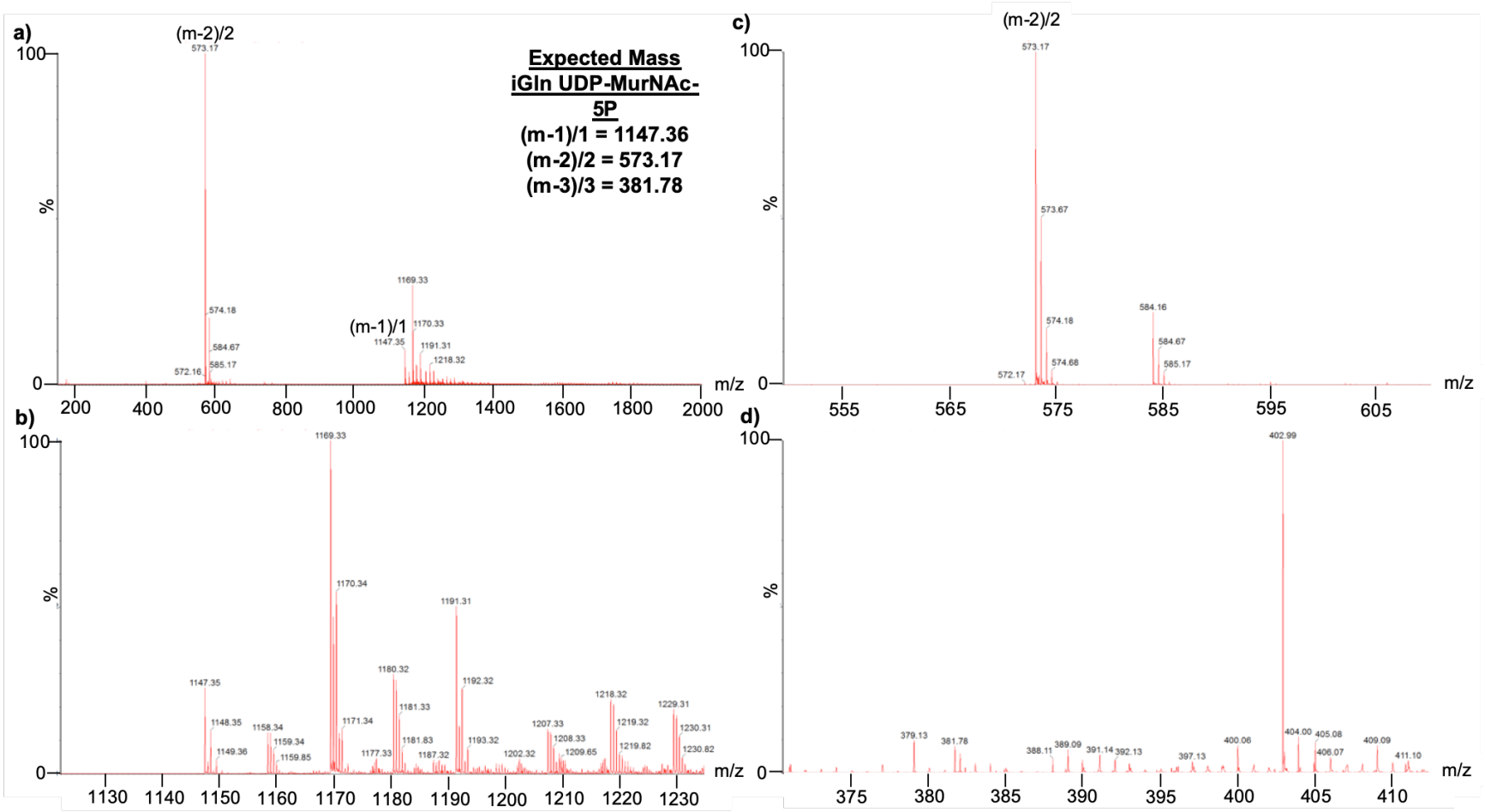
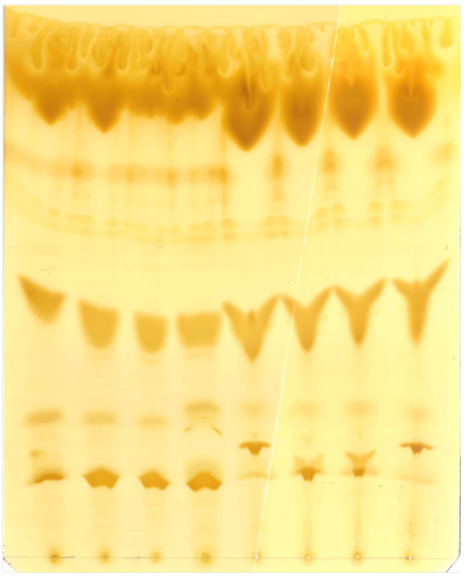



**Figure 4.20: Negative ion ES-MS of iGln UDP-MurNAc-5P** a) Full spectra, b) Enlarged spectra showing the series corresponding to (m-1)/1 (1147.36) c) Enlarged spectra showing the series corresponding to (m-2)/2 (573.17) d) Enlarged spectra showing the series corresponding to (m-3)/3 (381.78).



| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Overnight Incubation | ✓ | ✓ | ✓ | ✓ | | | | |
| Moenomycin | | | ✓ | ✓ | | | ✓ | ✓ |
| 2 x [iGln UDP-5P] | | ✓ | | ✓ | | ✓ | | ✓ |

**Figure 4.21: TLC of iGln Lipid II(Lys) synthesis optimisation**. Standard incubation (Lane 5) were modified with variations on length of incubation (3 hours vs overnight), addition of moenomycin or 2 x concentration of iGln UDP-Mur*N*Ac-5P as indicated by the table.

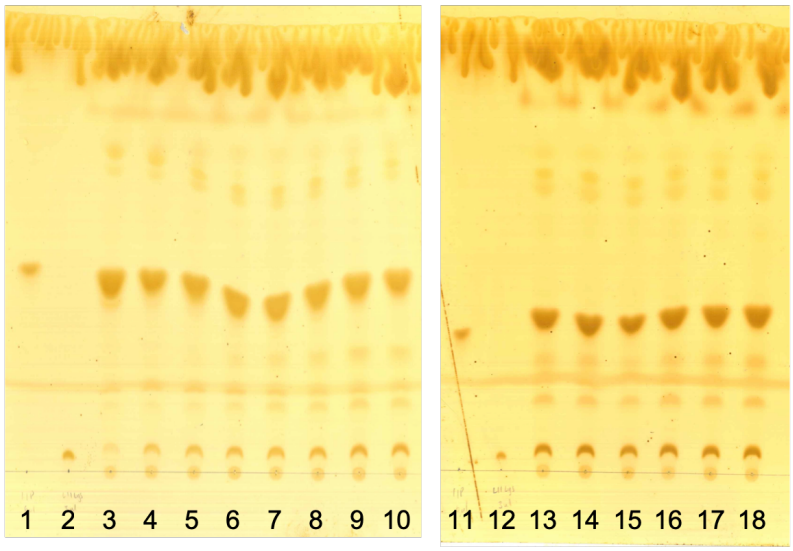



**Figure 4.22: TLC of iGln Lipid II(Lys) time course.** Lane 1 and 11 - Undecaprenyl pyrophosphate, Lane 2 and 12 - Lipid II(Lys) , Lanes 3 to 10 - hourly samples at T = 0 - 7 hours, Lanes 13 to 18 - hourly samples at T = 8 - 12 hours, Lane 18 - sample at T = 13.5

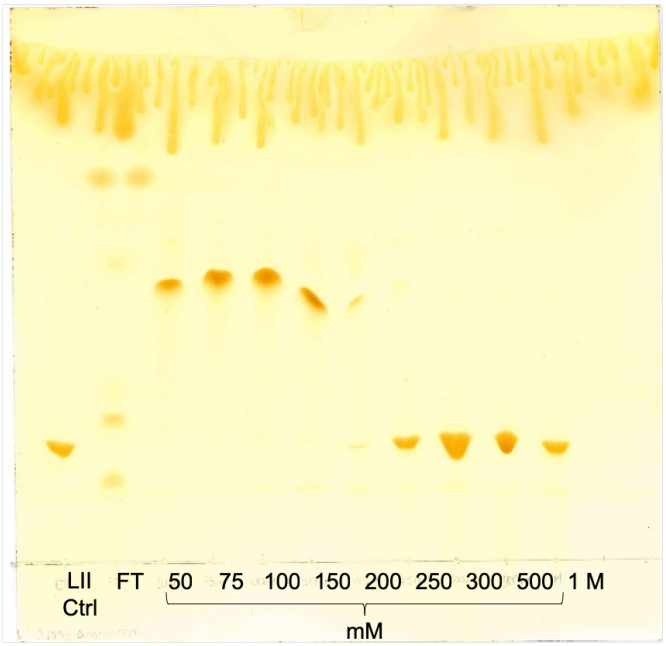
the product was eluted in 2:3:1 chloroform:methanol:ammonium bicarbonate when the stock concentration of ammonium bicarbonate was between 250 mM and 1 M. Since iGln Lipid II(Lys) is less negatively charged than Lipid II(Lys) it was found to bind the DEAE sephacryl more weakly, resulting in elution at lower concentrations of ammonium bicarbonate (i.e. $<$100 mM). Since the undecaprenyl-pyrophosphate also elutes between 50 mM and 100 mM, additional washes with small increments of ammonium bicarbonate were used to improved the purity of iGln Lipid II(Lys). The additional washes resulted in the separation of the undeccaprenyl-pyrophosphate and the iGln Lipid II(Lys) with them eluting at 50 - 75 mM and 80 - 100 mM respectively. TLC of final Lipid II(Lys) and iGln Lipid II(Lys) elutions are shown in Figure 4.23. The final yield of Lipid II(Lys) and iGln Lipid II(Lys) were 4 mg and 0.1 mg respectively.

Negative ion electrospray mass spectrometry was used to confirm the presence of Lipid II(Lys) (Figure 4.24) and iGln Lipid II(Lys) (Figure 4.25) following synthesis and purification. Some additional high intensity peaks which do not correspond to the m/z of Lipid II or iGln Lipid II are present and may be due to ammonium or sodium ion adducts, or solvent contamination. However it is important to note that the intensity of the peak does not necessarily reflect the abundance of that compound in the sample. Furthermore, this data is consistent with Lloyd *et al.* (2008) which confirmed the presence of Lipid II by fragmentation.

## 4.6 Preparation of pure *in vitro* transcribed tRNA isoacceptors and *S. pneumoniae (159)* crude tRNA

Sanderson and Uhlenbeck (2007) demonstrated that the presence of a G in either of the 51-63 bp position of tRNA conferred stability with EF-Tu. $tRNA^{Ser2}$ has a G at position 51 whilst $tRNA^{Ser1}$ and $tRNA^{Ser3}$ do not have a G at this position. In *S. pneumoniae* D39, the codons which base pair with the anticodon of $tRNA^{Ser1}$, $tRNA^{Ser2}$ and $tRNA^{Ser3}$ appear in 0.5, 0.81 and 1.45 % of the total genome. The

a)




b)

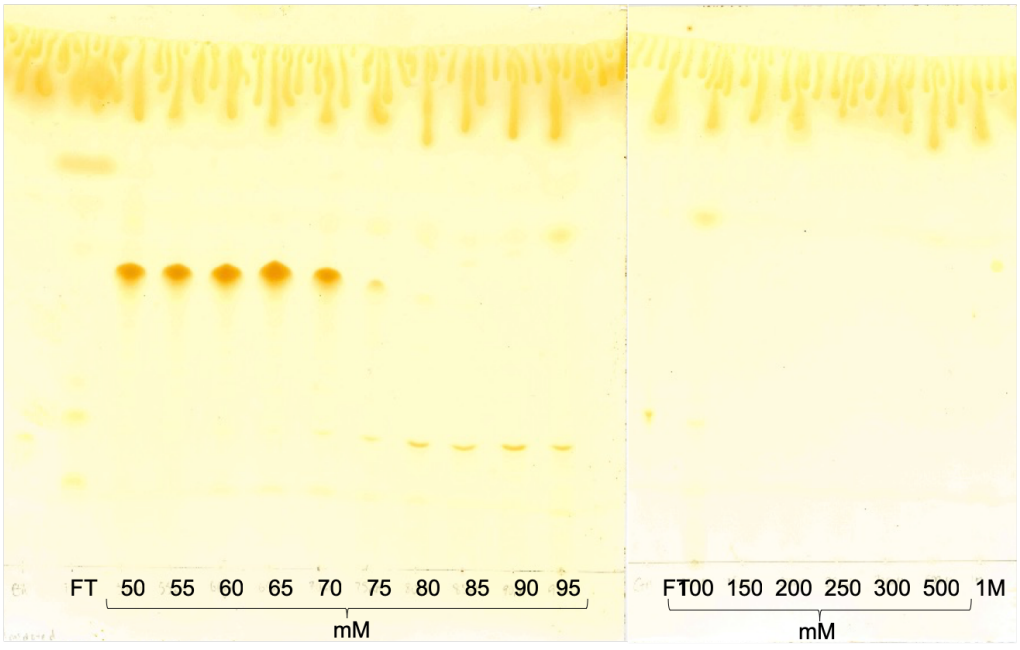



**Figure 4.23: TLC of Lipid II(Lys) and iGln-Lipid II(Lys) purifications.** a) Standard Purification of Lipid II(Lys) b) Optimised purification of iGln Lipid II(Lys).

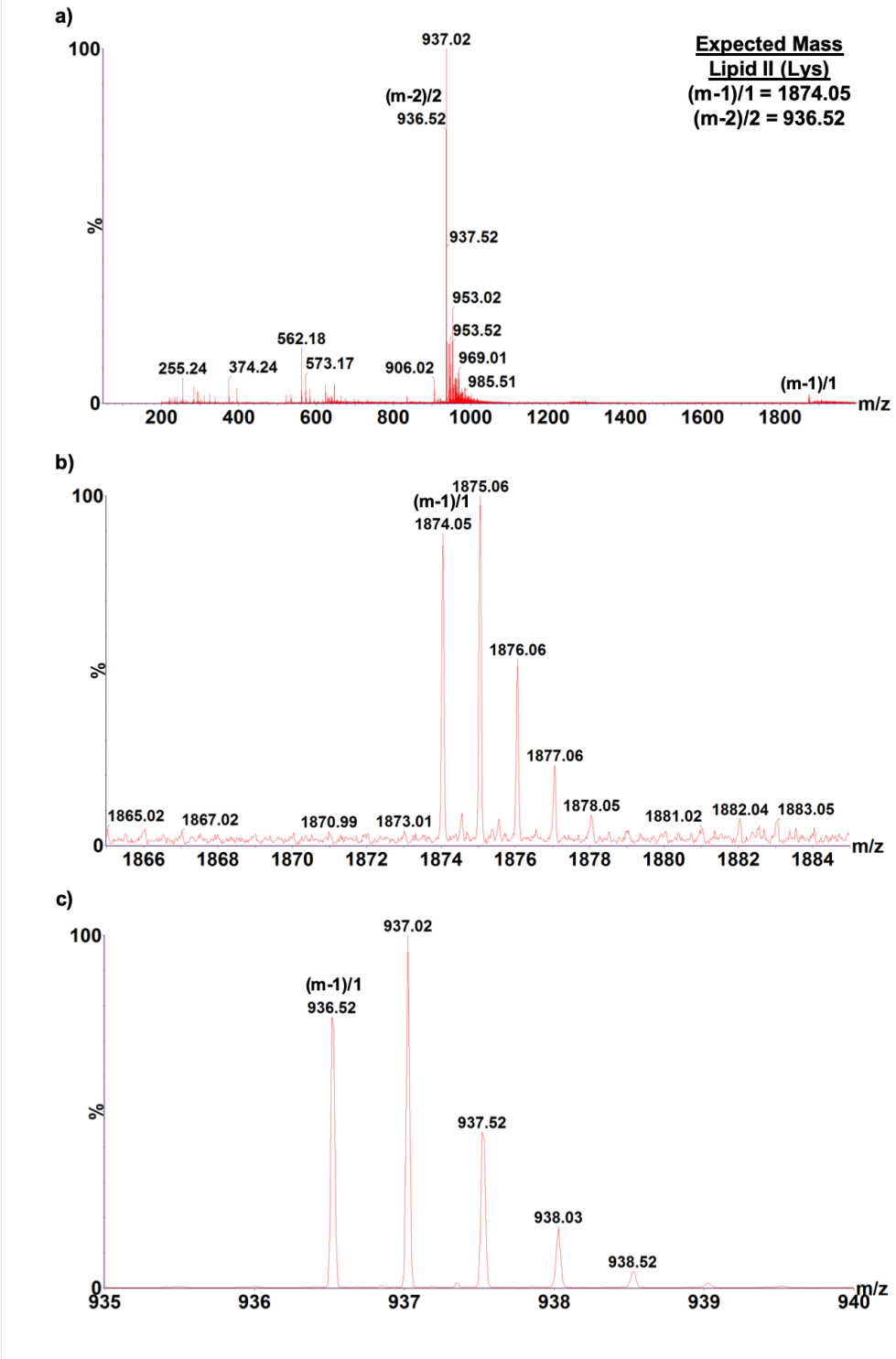



**Figure 4.24: Negative ion ES-MS of Lipid II(Lys)** a) Full spectra with *m/z* range 0 -2000, b) Enlarged spectra showing the series corresponding to(m-1)/1 (1874.05) c) Enlarged spectra showing the series corresponding to (m-2)/2 (936.52)

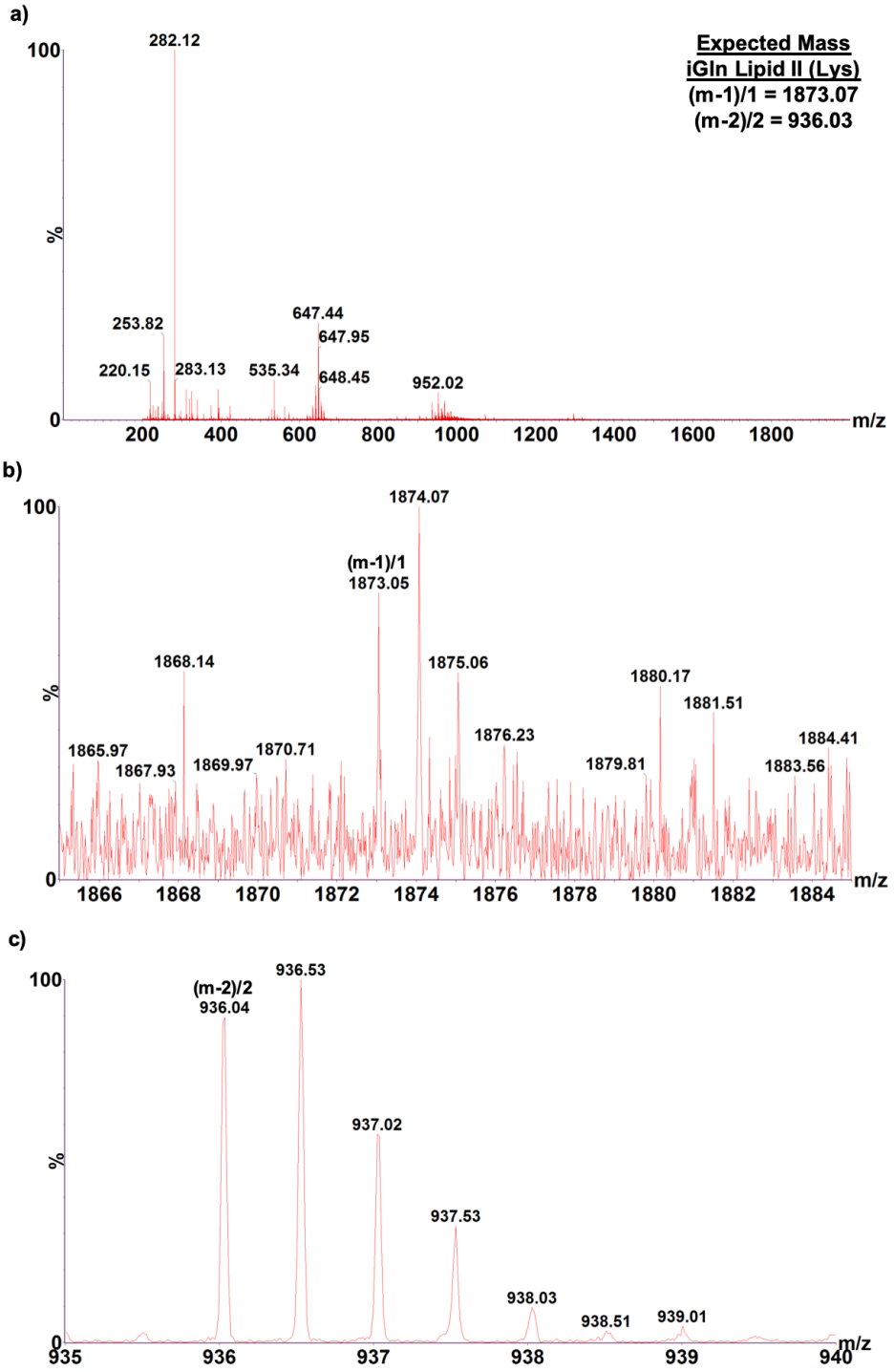



**Figure 4.25: Negative ion ES-MS of iGln Lipid II(Lys)** a) Full spectra with *m/z* range 0 -2000, b) Enlarged spectra showing the series corresponding to (m-1)/1 (1873.07) c) Enlarged spectra showing the series corresponding to (m-2)/2 (936.03)

availability of different tRNA isoacceptors for non-translational functions would depend upon a number of factors including tRNA isoacceptor expression levels, affinity for tRNA by aaRS, stability of interaction with EF-Tu and also codon usage within the organism (and therefore requirement in protein synthesis). In the absence of $H_2O_2$ MurM continues to synthesise branched PG, and so likely obtains cognately charged tRNAs which are surplus to protein synthesis. It is therefore possible that MurM may demonstrate bias towards specific isoacceptors. As such characterising kinetics of MurM with different isoacceptors would be useful for experiments going forward.

### 4.6.1 Alignment of tRNA isoacceptors

Isoacceptors of $tRNA^{Ala}$, $tRNA^{Ser}$ and $tRNA^{Thr}$ from *S. pneumoniae* R6 were aligned using GtRNAdb (Figure 4.26) (Chan and Lowe, 2009). This indicated that whilst $tRNA^{Ala}$ has only one isoacceptor, $tRNA^{Ser}$ has three isoacceptors and $tRNA^{Thr}$ has two isoacceptors. The tRNA isoacceptors present in the genome are able to supply all codons with amino acid, even in the absence of an exact codon-anticodon match, due to the presence of the wobble position. Further analysis of the $tRNA^{Ser}$(GGA), $tRNA^{Ser}$(GCU) and $tRNA^{Ser}$(UGA) isoacceptors identified that base pairs at positions 51:63 are U-C, C-U and G-C respectively, whilst $tRNA^{Thr}$(GGT) and $tRNA^{Ser}$(TGT) have base pairs C-G and G-T respectively. This suggests that $tRNA^{Ser}$(GGA) and $tRNA^{Ser}$(GCU) are less EF-Tu stable than $tRNA^{Ser}$(UGA) and so may be good candidates as high affinity MurM substrates.

$tRNA^{Ser}$(GGA), $tRNA^{Ser}$(GCU), $tRNA^{Ser}$(UGA), $tRNA^{Thr}$(GGU) and $tRNA^{Thr}$(UGU) from hereon in will be referred to as $tRNA^{Ser1}$, $tRNA^{Ser2}$, $tRNA^{Ser3}$, $tRNA^{Thr1}$ and $tRNA^{Thr2}$ respectively.

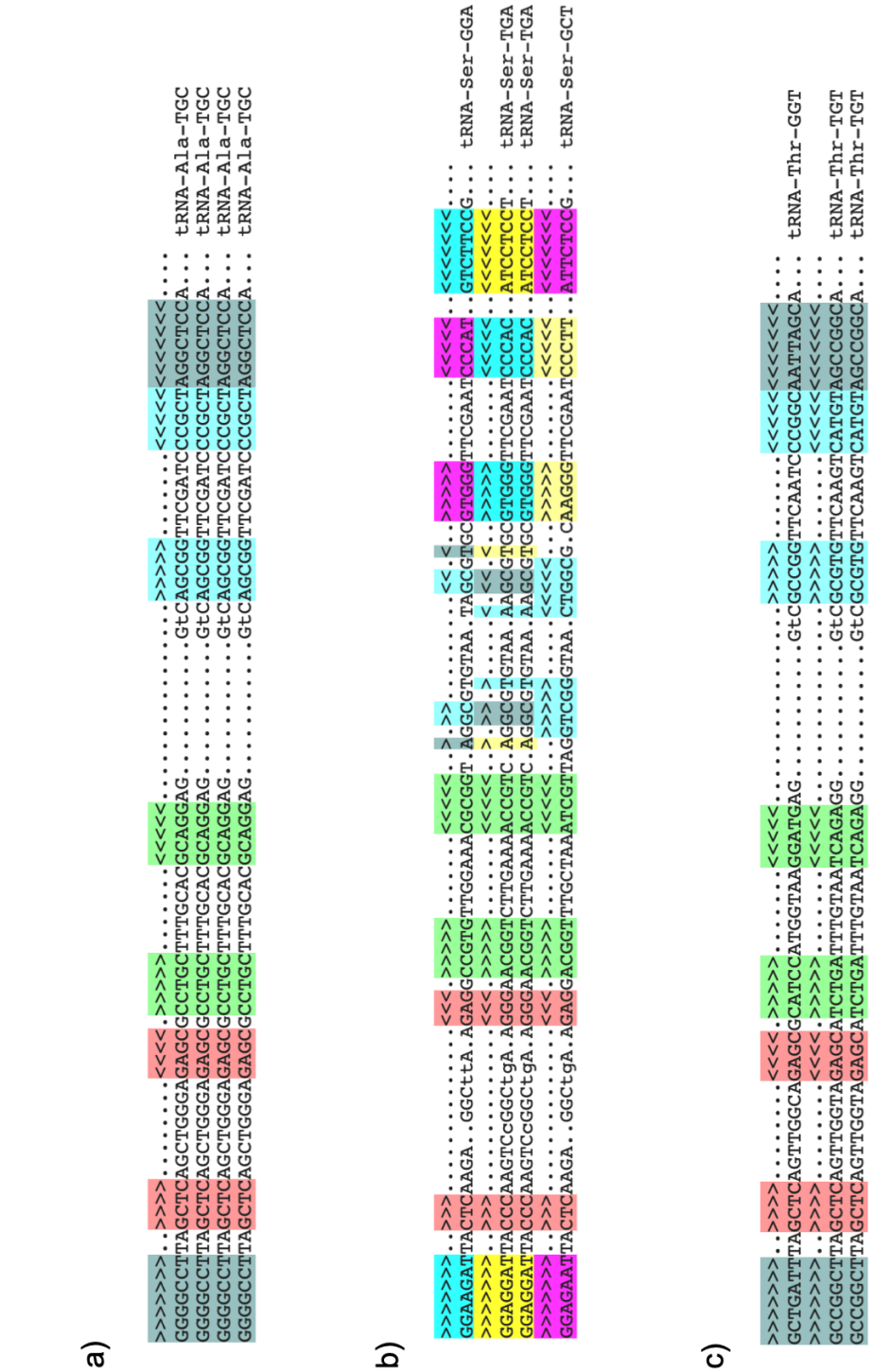



**Figure 4.26: tRNA gene alignments showing isoacceptors present in the *S. pneumoniae* (R6) genome.** a) $tRNA^{Ala}$, b) $tRNA^{Ser}$ and c) $tRNA^{Thr}$. Alignments conducted using GtRNAdb.

### 4.6.2 Purification of tRNA

*S. pneumoniae (159)* was grown for extraction of crude tRNA. Initially growth was conducted in the presence of whole blood and the cell yield was high. However, the whole blood would contain tRNA, which would cross-contaminate the *S. pneumoniae (159)* tRNA and so instead *S. pneumoniae (159)* was grown in the presence of purified bovine catalase (to a lower cell density). *In vitro* transcribed tRNA isoacceptors were purified as detailed in Section 2.8.

Both crude tRNA preparation and *in vitro* transcribed tRNAs can be seen in Figure 4.27. *S. pneumoniae* crude tRNA possesses a similar electrophoretic profile to *E. coli* crude tRNA, with a high density of bands corresponding to ∼75-100 bp. All bands of purified tRNA appear to run higher than the DNA marker of a corresponding size, this may be due to the tertiary structure present in tRNA. The pure *in vitro* transcribed tRNAs form single strong bands with few contaminants, however the isoacceptors did not follow the expected size based pattern on the gel ($tRNA^{Ala4}$ is 76 bp, $tRNA^{Ser1}$ and $tRNA^{Ser2}$ are 91 bp, $tRNA^{Ser3}$ is 93 bp and $tRNA^{Thr1}$ and $tRNA^{Thr2}$ are 76 bp). In addition, the band of $tRNA^{Thr2}$ shows smearing rather than a single band, smearing was observed in all *in vitro* transcription reactions for $tRNA^{Thr2}$ and was not effected by heating of tRNA samples prior to electrophoresis.

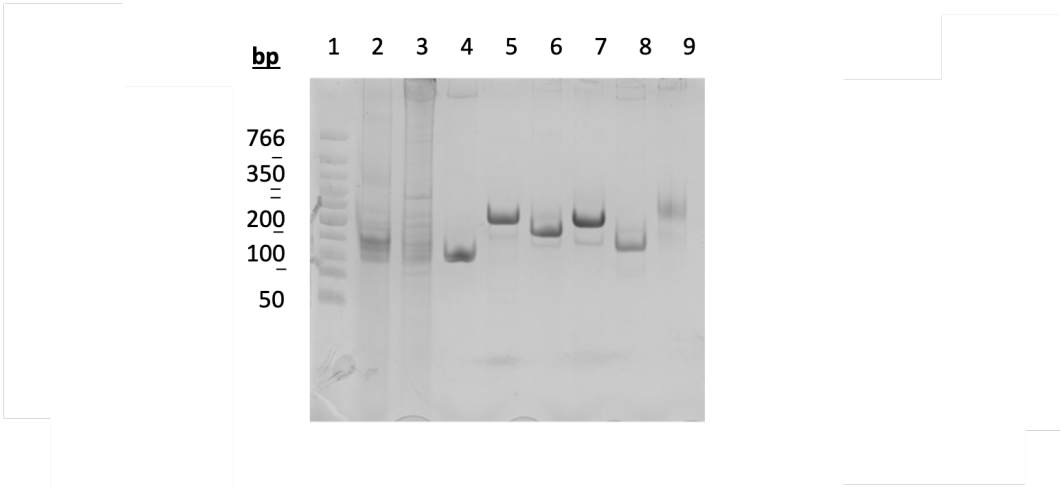



**Figure 4.27: Purified tRNA summary gel.** Lane 1 - NEB Low Molecular Weight DNA ladder, Lane 2 - 10 µg *E. coli* crude tRNA, Lane 3 - 10 µg *S. pneumoniae (159)* crude tRNA, Lane 4 - 5 µg pure $tRNA^{Ala}$, Lane 5 - 5 µg pure $tRNA^{Ser1}$, Lane 6 - 5 µg pure $tRNA^{Ser2}$, Lane 7 - 5 µg pure $tRNA^{Ser3}$, Lane 8 - 5 µg pure $tRNA^{Thr1}$, Lane 9 - 5 µg pure $tRNA^{Thr2}$. 7 M urea 12 % polyacrylamide gel.

## 4.7 Mass spectrometry MurM activity assays

### 4.7.1 Mass spectrometry of MurM assays

Mass spectrometry was used previously to demonstrate the addition of L-Ala or L-Ser to the third position lysine of the Lipid II(Lys) by $MurM_{159}$ (Lloyd *et al.*, 2008). These experiments were repeated with Lipid II(Lys) in order to confirm the activity of $gMurM_{159}$, and branching of the Lipid II(Lys) with both L-Ala and L-Ser. Products from reactions were identified by mass spectrometry.

Table 4.1 shows a summary of reactions conducted, the doubly charged expected mass (if the reaction was positive) and the observed mass from each sample.

The activity of $gMurM_{159}$ was confirmed by identification of an ion, with an *m/z* of 972.03 which corresponds to doubly charged species of Lipid II(Lys)-Ala in reaction 1. Reaction 2 is the negative control, the ion was absent from this spectra, instead an ion with an *m/z* of 936.52 which corresponds to a doubly charged unmodified Lipid II(Lys) was present (Figure 4.28). Reaction 3 indicated that in contrast to findings published by Lloyd *et al.* (2008) $gMurM_{159}$ cannot append L-Ser, charged by

SerRS onto $tRNA^{Ser}$, onto the Lipid II(Lys). Reactions 4-8 investigated the ability of $gMurM_{159}$ to utilise mis-acylated tRNAs (seryl-$tRNA^{Ala}$ or seryl-$tRNA^{Thr}$). Despite previous observations these results suggest that under these conditions $gMurM_{159}$ is unable to append L-Ser to Lipid II(Lys) from mis-aminoacylated tRNAs. The activity of SerRS was confirmed using a radiolabelled assay (Section 4.8) which

| **Reaction No.** | **Reaction components** | **Expected mass ((m-2)/2)** | **Observed mass ((m-2)/2)** |
|---|---|---|---|
| 1 | AlaRS, L-Ala, Lipid II(Lys), $gMurM_{159}$ | 972.042 | 972.029 |
| 2 | AlaRS, L-Ala, Lipid II (Lys) | 936.522 | 936.524 |
| 3 | SerRS, L-Ser, Lipid II(Lys), $gMurM_{159}$ | 980.037 | 936.52 |
| 4 | AlaRS, L-Ser, Lipid II(Lys), $gMurM_{159}$ | 980.037 | 936.52 |
| 5 | AlaRS(mutant), L-Ser, Lipid II(Lys), $gMurM_{159}$ | 980.037 | 936.52 |
| 6 | AlaRS($H_2O_2$ treated), L-Ser, Lipid II(Lys), $gMurM_{159}$ | 980.037 | 936.52 |
| 7 | ThrRS, L-Ser, Lipid II(Lys), $gMurM_{159}$ | 980.037 | 936.52 |
| 8 | ThrRS($H_2O_2$ treated), L-Ser, Lipid II(Lys), $gMurM_{159}$ | 980.037 | 936.52 |

**Table 4.1: Summary of results from $gMurM_{159}$ activity assays**. Negative ion mass spectrometry of Lipid II purified from $gMurM_{159}$ activity assays. Reaction components, predicted and observed mass of doubly charged species are presented. The expected (m-2)/2 values for Lipid II(Lys), Lipid II(Lys)-Ala and Lipid II(Lys)-Ser were 936.52, 972.04 and 980.04 respectively.

eliminated SerRS activity as a possible issue in these assays. The ability of MurM to utilise seryl-tRNA$^{Ser}$ and mis-acylated tRNAs can also be investigated using the radiolabelled MurM activity assay.

### 4.7.2 Tandem mass spectrometry characterising the Lipid II(Lys) product

Collision Induced Dissociation (CID) combined with positive ion tandem mass spectrometry (MS/MS) of the singly charged molecular ion of Lipid II(Lys)-Ala (1947.1/1), was used to confirm the exact location of the apparent addition of L-Ala to Lipid II(Lys) by gMurM$_{159}$ in reaction 1 (Table 4.1).

Figure 4.29 shows the annotated CID positive ion MS/MS spectra at Collision Energy of 75 kV, fragment ions were singly charged. Table 4.3 shows that peak 28 has a mass of an additional 71.0004 when compared to peak 27. The additional mass of 71.0004 corresponds to the exact mass expected by the condensation of alanine with the $\epsilon$-amino group of the third position stem peptide lysine. The mass of peak 27 in the negative control (-gMurM$_{159}$) corresponds to the predicted mass of GlcNAcMurNAc-AEKAA. This result indicates that the addition of a single L-Ala occurred in the test sample (+gMurM$_{159}$) but not in the negative control (-gMurM$_{159}$).

Peaks 6 and 7 have observed *m/z* values corresponding to K**A**AA and EK**A**A or AEK**A** respectively and can be observed in the +gMurM$_{159}$ sample but are absent from the -gMurM$_{159}$ spectra, instead peaks 4 and 5 were observed at *m/z* values corresponding to the linear fragments (without an additional L-Ala appended to the third position lysine). The masses of peaks 6 and 7 are consistent with the addition of L-Ala to the $\epsilon$-amino group of the third position lysine of the Lipid II(Lys) pentapeptide.

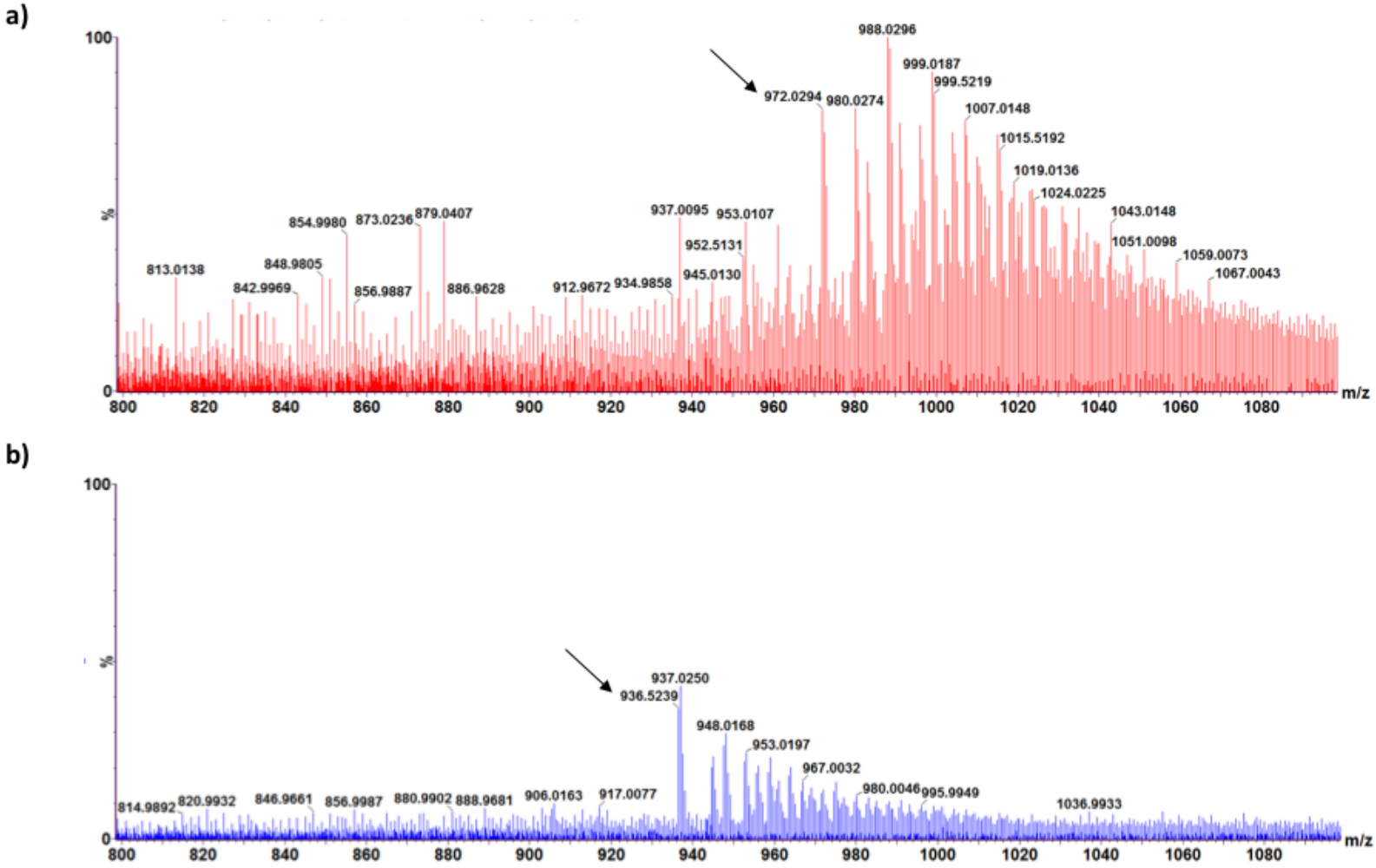



**Figure 4.28: Negative ion mass spectra for MurM activity assays.** a) in the presence of $gMurM_{159}$ or b) in the absence of $gMurM_{159}$. Arrows indicate the major peak associated with either the modified or unmodified Lipid II(Lys) species.

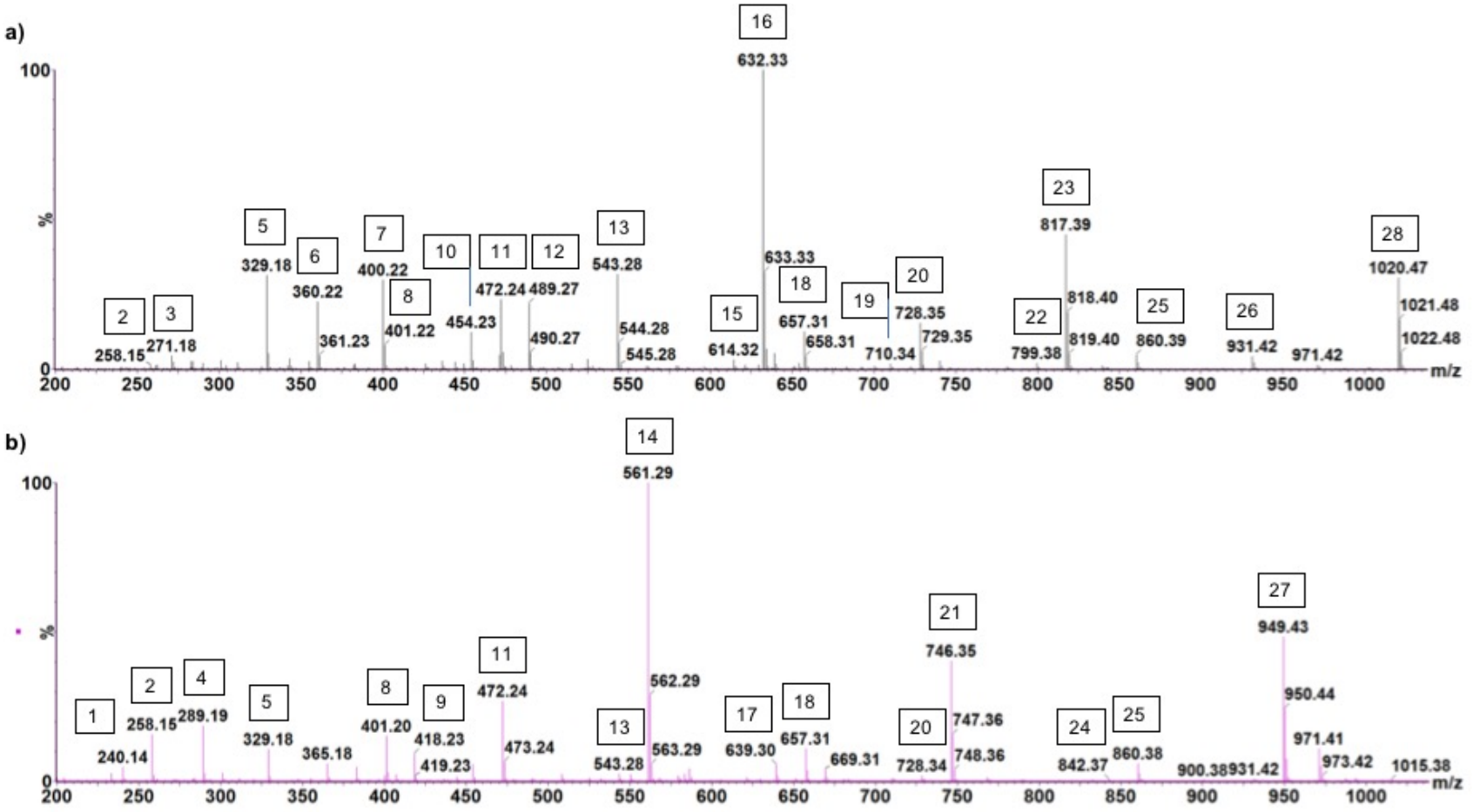



**Figure 4.29: Positive ion mass spectra from Lipid II(Lys) fragmentation.** Collision Induced Dissociation in combination with positive ion tandem mass spectrometry (MS/MS) fragmentation spectra, for purified Lipid II(Lys) following a MurM activity assay in the a) presence of $gMurM_{159}$ or b) absence of $gMurM_{159}$. Collision energy of 75 kV provided the clearest range of fragments over the length of the Lipid II(Lys) species from the MurNAc to the terminal D-ala. The observed peak masses were annotated with unique numbers and assigned to potential different fragments of corresponding expected masses (detailed in Table 4.3).

| **Peak No.** | **Possible Fragments** | ***m/z*** |
|---|---|---|
| 1 | MurNAc (minus $H_2O$) | 240.0872 |
| 2 | EK | 258.1454 |
| 3 | KAA | 289.1876 |
| 4 | EKA or AEK or EK**A** | 329.1825 |
| 5 | AKAA | 360.2247 |
| 6 | EK**A**A or AEK**A** | 400.2196 |
| 7 | Lactyl AEK | 401.2036 |
| 8 | EKAA | 418.2302 |
| 9 | Lactyl AEK**A** (minus $H_2O$) or Lactyl AEKA (minus $H_2O$) | 454.2302 |
| 10 | Lactyl AEK**A** or Lactyl AEKA | 472.2407 |
| 11 | EK**A**AA or AEKAA | 489.2673 |
| 12 | Lactyl AEK**A**A or Lactyl AEKAA (minus $H_2O$) | 543.2778 |
| 13 | Lactyl AEKAA | 561.2884 |
| 14 | Lactyl AEK**A**AA (minus $H_2O$) | 614.31497 |
| 15 | Lactyl AEK**A**AA | 632.3255 |
| 16 | MurNAc AEKA (minus $H_2O$) | 639.2990 |
| 17 | MurNAc AEKA | 657.3096 |
| 18 | MurNAc AEK**A**A (minus $H_2O$) | 710.3361 |
| 19 | MurNAc AEKAA (minus $H_2O$) or AEK**A**A | 728.3467 |
| 20 | MurNAc AEKAA | 746.3572 |
| 21 | MurNAc AEK**A**AA (minus $H_2O$) | 799.3838 |
| 22 | MurNAc AEK**A**AA | 817.3944 |
| 23 | GlcNAcMurNAc AEKA (minus $H_2O$) | 842.3784 |

| | | |
|---|---|---|
| 24 | GlcNAcMurNAc AEK**A** or GlcNAcMurNAc AEKA | 860.3889 |
| 25 | GlcNAcMurNAc AEK**A**A or GlcNAcMurNAc AEKAA (minus $H_2O$) | 931.4260 |
| 26 | GlcNAcMurNAc AEKAA | 949.4366 |
| 27 | GlcNAcMurNAc AEK**A**AA | 1020.4737 |

**Table 4.3: Fragments assigned to peaks in Figure 4.29.** Peak number, the assigned possible fragments and the associated expected *m/z*. All amino acids are shown sequentially, those not in bold are part of the main pentapeptide chain, whilst any amino acid shown in bold font, is appended to the $\epsilon$-amino group of the third position lysine of the 5P stem. Note: single letter code for amino acids used in this table.

## 4.8 Preparation for radiolabelled MurM activity assays

Radioactive assays which follow the transfer of [$^3$H]aa from [$^3$H]aa-tRNA$^{aa}$ to Lipid II(Lys) have previously been used to monitor the activity of MurM (Lloyd *et al.*, 2008; Shepherd, 2011).

Radioactive MurM activity assays with both Lipid II(Lys) and iGln Lipid II(Lys) as well as different tRNA isoacceptors will provide information on the substrate specificity of MurM. MurM kinetics with correctly acylated tRNA and mis-aminoacylated tRNA could also be investigated using this assay. Answering these questions will help elucidate any link between MurM and protein synthesis, especially under oxidative stress where $H_2O_2$ levels are high and mis-aminoacylation of tRNA increases.

Initially the ability of AlaRS, SerRS and ThrRS to aminoacylate cognate *S. pneumoniae* tRNAs, creating substrates for MurM was demonstrated.

The transfer of free [$^3$H]aa to tRNA by aaRS was followed over 90 mins by precipitation of tRNA in 10 % TCA (w/v) and scintillation counting was used

to measure [$^3$H]aa-tRNA$^{aa}$. A negative control (-tRNA), at corresponding time points, and an absolute background (scintillant only) reading were deducted from the measured counts per minute (cpm) for every sample. Figure 4.30 shows that AlaRS, SerRS and ThrRS are capable of charging *S. pneumoniae (159)* crude tRNA with [$^3$H]aa. The extent of acylation varies between AlaRS, SerRS and ThrRS and likely reflects the total amount of tRNA$^{Ala}$, tRNA$^{Ser}$ and tRNA$^{Thr}$ in the crude tRNA. In all cases the the vast majority of aminoacylation has occurred in the first minute with some gradual increase occurring over the remaining 90 minute period.

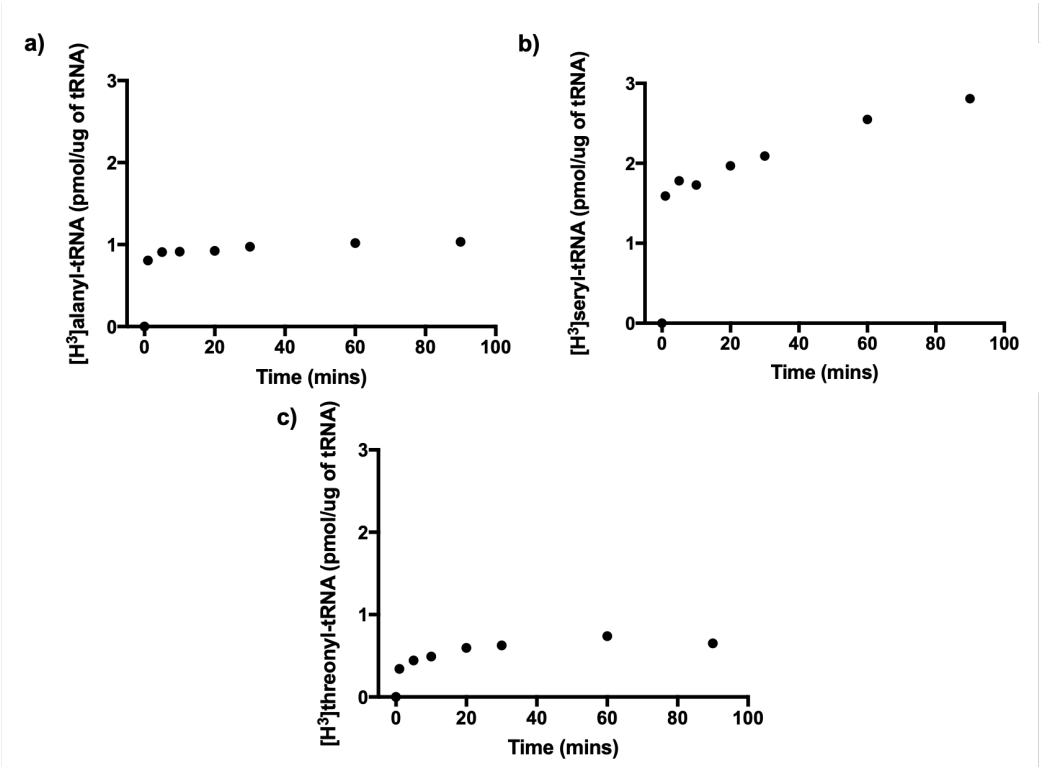



**Figure 4.30: Aminoacylation of *S. pneumoniae (159)* crude tRNA by AlaRS, SerRS and ThrRS in the presence of cognate amino acids** a) charging of AlaRS in the presence of L-alanine b) charging of SerRS in the presence of L-serine c) charging of ThrRS in the presence of L-threonine. The specific activity of [$^3$H]L-alanine, [$^3$H]L-serine and [$^3$H]L-threonine were calculated to be 128.5 cpm/pmol, 46.8 cpm/pmol and 69.7 cpm/pmol respectively.

Aminoacylation of the *in vitro* transcribed tRNAs with their cognate amino acids were investigated (Figure 4.31). From this, it is clear that the level of aminoacylation was highest for tRNA$^{Ser}$, where all three tRNA$^{Ser}$ isoacceptors acylated well. tRNA$^{Ala}$ did undergo aminoacylation but the final yield was a lot lower, and for

$tRNA^{Thr}$ isoacceptors very little aminoacylation occurred. The reason for low levels of aminoacylation with $tRNA^{Ala}$ and $tRNA^{Thr}$ isoacceptors may indicate an issue with the *in vitro* transcribed tRNA, or may more be reflective of the activity of AlaRS and ThrRS. If the aaRS activity is low, then spontaneous deacylation may be limiting. It would therefore be useful to repeat this experiment with higher concentrations of AlaRS and ThrRS.

Since *S. pneumoniae (159)* tRNA is limited and *E. coli* tRNA is readily available from Sigma Aldrich, both preparative charging and MurM assays were conducted with *E. coli* crude tRNA as a preliminary test. These assays were to ensure that the assay set up was appropriate, to gain a better idea of aminoacylated tRNA yields, and to establish appropriate working concentrations/ cpm readings for MurM assays.

Preparative charging of 0.38 $mg.mL^{-1}$ *E. coli* tRNA was conducted with AlaRS and [$^3$H]L-alanine, the sample was incubated at 37 °C for 1 hour. The final yield of [$^3$H]alanyl-$tRNA^{Ala}$ was 50 µl of 10 $\mu$M. The losses incurred during purification of the [$^3$H]alanyl-$tRNA^{Ala}$ were $\sim$25 %.

An end-point radioactive MurM assay was used to test; the functionality of the assay, the [$^3$H]alanyl-$tRNA^{Ala}$ from the previous preparative charging experiment, and to compare the ability of MurM to use Lipid II(Lys) iGln Lipid II(Lys) as substrates. Table **??** shows the cpm for each test reaction and all the negative controls. The cpm were well within an easily detectable range and were considerably higher than readings from all the negative controls. This demonstrated that tRNA from preparative charging provided a suitable substrate, that $gMurM_{159}$ was active and that the radioactive MurM assay worked as expected. In addition this assay demonstrated that $gMurM_{159}$ was capable of utilising both Lipid II(Lys) and iGln Lipid II(Lys) substrates. However, it does not provide any information regarding the rates of each reaction; kinetic experiments would be required in order to determine the relative catalytic efficiency of these two substrates.

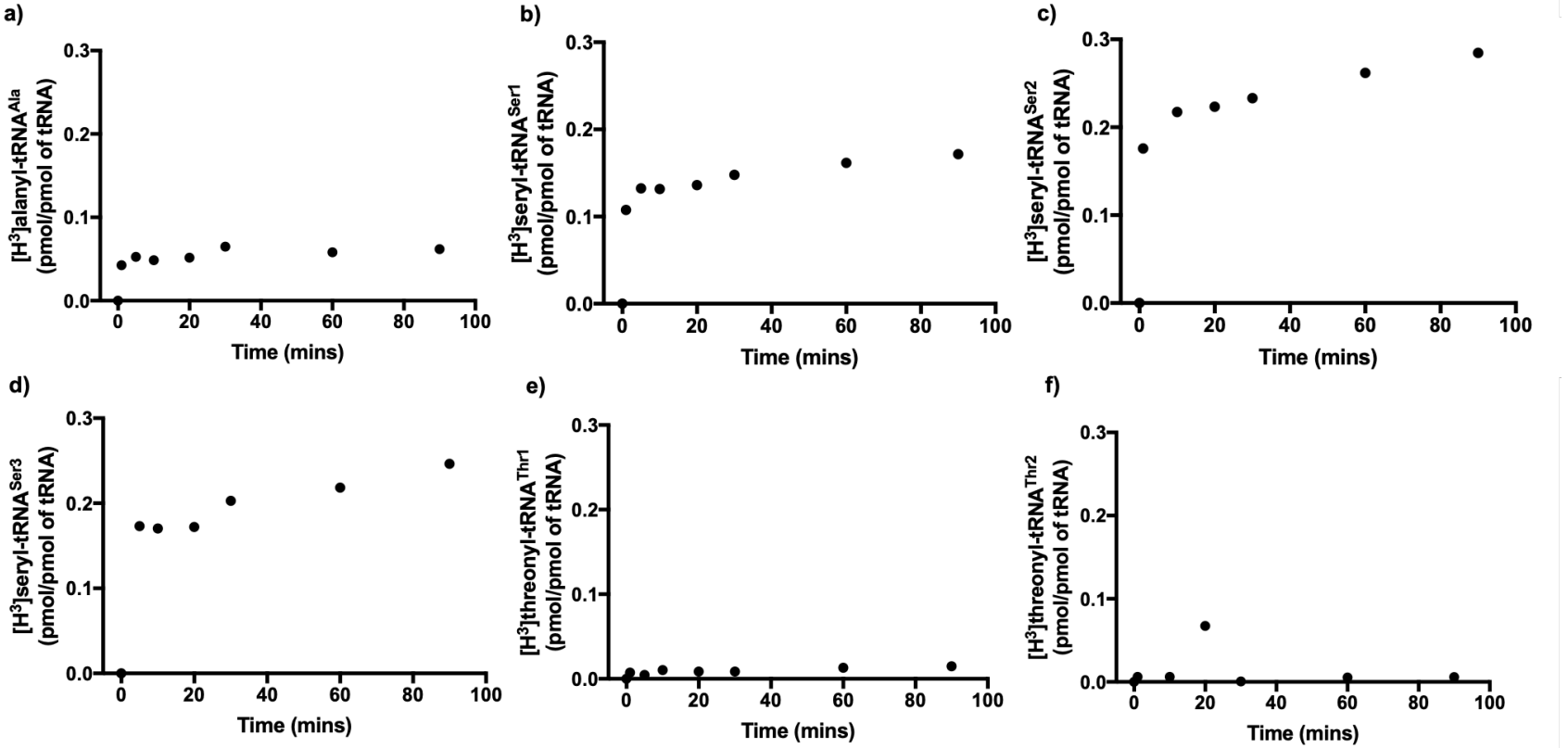



**Figure 4.31: Aminoacylation of *in vitro* transcribed pure tRNA isoacceptors by AlaRS, SerRS and ThrRS in the presence of cognate amino acids**
a) aminoacylation of $tRNA^{Ala}$ by AlaRS in the presence of L-Ala, b) aminoacylation of $tRNA^{Ser1}$ by SerRS in the presence of L-Ser, c) aminoacylation of $tRNA^{Ser2}$ by SerRS in the presence of L-Ser, d) aminoacylation of $tRNA^{Ser3}$ by SerRS in the presence of L-Ser, e) aminoacylation of $tRNA^{Thr1}$ by ThrRS in the presence of L-Thr f) aminoacylation of $tRNA^{Thr2}$ by ThrRS in the presence of L-Thr. The specific activity of [$^3$H]L-Ala, [$^3$H]L-Ser and [$^3$H]L-Thr were calculated to be 70.4 cpm/pmol, 63.1 cpm/pmol and 70.8 cpm/pmol respectively.

| Reaction | cpm of Lipid II(Lys) * | cpm of iGln Lipid II(Lys) |
|---|---|---|
| Complete reaction | 2213 | 1796 |
| -gMurM$_{159}$ | 206 | 162 |
| +RNase | 259 | 197 |
| -Lipid II(Lys) | 176 | 127 |

**Table 4.4: Summary of end-point MurM radioactivity assay results.** gMurM$_{159}$ and three negative controls; -gMurM$_{159}$, +RNase and -Lipid II(Lys) were incubated in a radioactivity assay at 37 °C for 1 hour. *Raw cpm data has been normalised to the second data set based on total radioactivity.

## 4.9 Discussion and conclusions

This chapter demonstrates significant progress in preparation for a thorough investigation into substrate specificity of gMurM$_{159}$. Initially, since *S. pneumoniae (159)* has a higher proportion of L-alanine in the dipeptide cross bridge and *S. pneumoniae (Pn16)* has a higher proportion of L-serine, it was intended that both gMurM$_{159}$ and gMurM$_{Pn16}$ would be investigated and compared. However, due to time constraints and the delays due to difficulties outlined in this chapter, only gMurM$_{159}$ was expressed, purified and tested.

Throughout this chapter, challenges in expression of gMurM$_{159}$ and gMurM$_{Pn16}$ have been overcome. The sequences of gMurM$_{159}$ and gMurM$_{Pn16}$ have been confirmed and full genome sequencing of *S. pneumoniae (159)* and *S. pneumoniae (Pn16)* has been conducted. Both MurM substrates (Lipid II(Lys) and tRNA) and variants thereof have been synthesised and purified for downstream experiments. Preliminary experiments have confirmed gMurM$_{159}$ activity and ability of aaRSs to generate aminoacylated tRNA.

The synthesis and purification of iGln Lipid II(Lys) was optimised, however the yield still remained low ($\sim$40-fold lower than Lipid II(Lys) ). Radioactive MurM end-point assays indicated that both Lipid II(Lys) and iGln Lipid II(Lys) are substrates for

MurM, however, further kinetic analysis is required.

*In vitro* transcription and purification of isoacceptors of $tRNA^{Ala}$, $tRNA^{Ser}$ and $tRNA^{Thr}$ was conducted. $tRNA^{Ala}$ is 76 bp, $tRNA^{Ser1}$ $tRNA^{Ser2}$ are 91 bp, $tRNA^{Ser3}$ is 93 bp and $tRNA^{Thr1}$ $tRNA^{Thr2}$ are 76 bp, however the migration on 7 M urea 12 %-PAGE did not reflect these values. The isoacceptors template DNA did follow the correct size pattern on a 3 % agarose gel, indicating that the apparent incorrect masses may be due to residual tertiary structure of different isoacceptors. However, it is also important to consider that these differences in expected size may indicate premature termination of tRNA during *in vitro* transcription which would cause issues in downstream experiments due to the loss of the 3' CCA, which is required for charging of tRNA by the aaRSs. As such it would be useful to confirm the identity of the tRNA products by mass spectrometry (Huang *et al.*, 2010). Heterogeneity of length or tertiary structure may also explain the smearing of $tRNA^{Thr2}$ in a 7 M urea 12 %-PAGE.

In contrast to Lloyd *et al.* (2008) $gMurM_{159}$ was only able to append L-Ala and not L-Ser onto Lipid II(Lys). One such explanation for this observation may be the experimental design and the kinetics of the enzymes involved. The total amount of tRNA used in these reactions was ∼7-fold lower than was used by Lloyd *et al.* (2008). Since in *S. pneumoniae* the total isoacceptors for $tRNA^{Ala}$ are ∼2.5-fold more abundant than $tRNA^{Ser}$ isoacceptors (Lloyd *et al.*, 2008), even in the presence of a 7-fold reduction in total tRNA, AlaRS may have been able to turnover $tRNA^{Ala}$ at a rate suitable to provide substrate for $gMurM_{159}$. In contrast, for SerRS, the low concentration of $tRNA^{Ser}$ may have resulted in a rate of aminoacylation slower than that of deacylation. If the rate of deacylation was too rapid for $gMurM_{159}$ to complete with there would have been no seryl-$tRNA^{Ser}$ substrate available for $gMurM_{159}$. Therefore, reactions with SerRS should be repeated with a higher concentration of total tRNA. Similarly, all reactions exploring mis-aminoacylation (Figure 4.1 reactions 6-10) should be repeated at a higher concentration of aaRS, tRNA and L-Ser, since the rate of mis-aminoacylation would be significantly slower than

cognate aminoacylation. The MurM assays therefore do not conclusively show that $gMurM_{159}$ is unable to utilise $tRNA^{Ser}$ as a substrate, but instead indicates that the experiments need to be repeated with more appropriate substrate concentrations and considerations to the kinetic parameters of each reaction. The radiolabelled MurM activity assays can be used to investigate the ability of MurM to utilise purified charged tRNA substrates directly, without the coupling of aaRS and MurM activity.

## 4.10 Future experiments

The future work for this chapter consists of large scale expression and purification of $gMurM_{Pn16}$, confirmation of activity and use in final kinetic experiments.

The identity of *in vitro* transcribed tRNAs may be confirmed by mass spectrometry as per Huang *et al.* (2010).

Mass spectrometric MurM activity assays will be optimised in order to demonstrate the addition of L-serine to Lipid II(Lys) by MurM from tRNAs charged with both cognate and non-cognate amino acids. In addition these assays will be used to demonstrate the addition of L-Ala and L-Ser to iGln Lipid II(Lys) by MurM from tRNAs charged with both cognate and non-cognate amino acids.

A number of sequential experiments will be required in order to investigate the hypothesis that MurM preferentially utilises mis-aminoacylated tRNAs in order to maintain the fidelity of protein synthesis. *S. pneumoniae* crude tRNA and *in vitro* transcribed tRNAs will be aminoacylated or mis-aminoacylated using the appropriate synthetase and conditions. These acylated tRNAs will then provide substrate for MurM assays. The kinetic parameters of MurM with Lipid II(Lys) and iGln Lipid II(Lys) will be investigated. Whilst maintaining a constant concentration of Lipid II(Lys) the effect of varying crude tRNA concentration will be investigated. These experiments will aim to establish the substrate specificity of MurM for crude

extract containing alanyl-tRNA$^{Ala}$, seryl-tRNA$^{Ser}$, seryl-tRNA$^{Ala}$ and seryl-tRNA$^{Thr}$. Subsequently different isoacceptors of tRNA$^{Ser}$ and tRNA$^{Thr}$ as substrates for gMurM$_{159}$ and gMurM$_{Pn16}$ will be investigated.

In addition, since other synthetases such as ProRS and LysRS are known to mis-aminoacylate tRNA with L-Ala and L-Ser (Beuning and Musier-Forsyth, 2000; Shepherd and Ibba, 2014), it would be prudent to investigate whether MurM also utilises these mis-aminoacylated tRNAs as substrates.

# Chapter 5

# Computational modelling of $MurM_{159}$

The work in the chapter was conducted in collaboration with Dr. Charo I. del Genio (School of Life Sciences, University of Warwick) for homology modeling, Prof. Rudo Roemer (Department of Physics, University of Warwick) for flexibility simulations and Prof. Syma Khalid (School of Chemistry, University of Southampton) for protein-membrane interaction simulations.

## 5.1 Introduction

Insight from protein structure is commonly used to assist structure-based drug design, elucidation of enzyme mechanisms and specificity of protein-ligand interactions. Additionally, structural knowledge can aid experimental design and focus future research, for example providing target residues for rational site directed mutagenesis (Smyth and Martin, 2000). The main experimental techniques to determine the structure of proteins, and protein:protein complexes, both in the presence and absence of ligands are X-ray crystallography, Nuclear Magnetic Resonance (NMR) and Electron Microscopy (EM).

X-ray crystallography is a widely used, powerful technique which can be used to help answer complex biological questions. Potential dynamic movements within the protein may be inferred or sometimes modelled in the presence of multiple different structures of the same protein (Acharya and Lloyd, 2005). Furthermore, crystal structures represent time and space averages of all molecules present within the crystal lattice, therefore, conformational variations between these can provide information about the flexibility of a region, which may or may not be important for function (Acharya and Lloyd, 2005). X-ray crystallography does however have its own limitations since crystallography only provides a snap-shot of a protein in a particular conformation; proteins are dynamic structures and the presence of substrates, co-factors or other proteins may dramatically alter their conformation. Furthermore, some structures occur as a direct consequence of the protein being in a crystalline state, and so it is important to consider the biological relevance of any structure (Acharya and Lloyd, 2005). Collection of X-ray crystallography data at ambient temperature causes radiation damage such that complete diffraction data sets must be collected from multiple crystals, as such it is common to cryogenically freeze crystals and collect data at liquid nitrogen temperature. However, it has been shown that cryogenic freezing can remodel conformational distributions of more than 35 % of side chains, eliminate packing defects necessary for functional motion and hinder identification of sites with the potential to allosterically modulate protein function. Therefore, cryogenically frozen and room-temperature data collection may be used in tandem to reveal motions crucial for catalysis, ligand binding and allosteric regulation which would otherwise be lost (Fraser *et al.*, 2011; Fischer *et al.*, 2015). Importantly, however, data collection at room temperature results in radiation damage such that complete diffraction data sets must be obtained from multiple crystals.

With the advancement of DNA sequencing techniques it is evident that it will not be possible to obtain structural information on all proteins of interest using the current experimental methods (Schwede, 2013). However, homology modelling, a

computational technique also known as comparative modelling can "be applied with reasonable accuracy to ten times more protein sequences than the number of experimentally determined protein structures" (Sanchez and Šali, 1997). In the absence of experimental 3D protein structures, homology models can direct and focus future experiments and assist analysis of existing experimental data.

Homology modelling consists of four steps: identification of related proteins with experimentally determined structures; sequence alignment of these and the target to identify a 'template' structure; build a model of the target protein based upon the structure; and sequence alignment of the template and finally evaluation of the model (Sanchez and Šali, 1997). A huge number of different software packages are available for the homology modelling of proteins, including The SWISS-MODEL workspace (Arnold *et al.*, 2006), SegMod (Levitt, 1992) and MODELLER (Eswar *et al.*, 2006). These examples represent three different methods used to build the model of the target protein; rigid-body assembly, modelling by segment matching and modelling by satisfaction of spatial restraints respectively (Sanchez and Šali, 1997). As a result of slightly different methods, the quality and accuracy of the models can vary significantly between different software/servers, however Schwede *et al.* (2009) acknowledge that even lower resolution models are useful for focused experimental design (e.g site directed mutagenesis) and for supporting experimental structure determination. Additionally the time taken for the result to be returned can vary between different software/servers, ranging from hours to weeks. Ultimately, however, the higher the sequence identity between template and target, the better the quality of the model.

Determining protein flexibility as well as identifying interactions with protein binding partners, ligands or membranes, can be important in understanding protein function and elucidating mechanisms of action. However, proteins have a high number of degrees or freedom, which can be challenging for computational modelling methods. Molecular Dynamics (MD) is a simulation method for looking at the physical movement of atoms of molecules within a structure. This type of

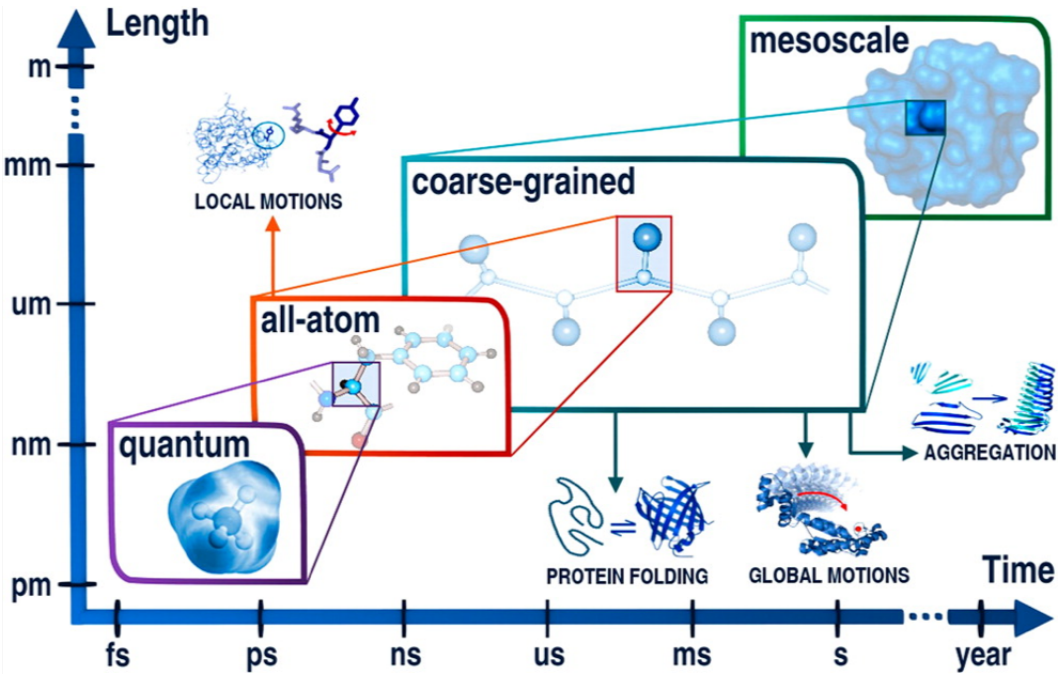



**Figure 5.1: Application ranges for molecular modeling at different resolutions: quantum, all-atom, coarse-grained, and mesoscale.** The plot shows approximate ranges of time scales and system sizes (lengths). The presented application ranges can be expanded by merging tools of different resolution into multiscale schemes. Figure sourced from Kmiecik *et al.* (2016).

atomistic representation results in most accurate and reliable results, however it can be time-consuming and computer intensive. Due to the limitations of available computing power, atomistic simulations can only be conducted for biological processes occurring in very short (picosecond to nanosecond) time-frames. Coarse-grained modelling is a method which has gained popularity and provides a means of circumventing these issues (Hospital *et al.*, 2015). The overall resolution (granularity) of the model is reduced such that instead of modelling individual atoms, atoms are grouped together into pseudo-atoms which can then be modelled as an overall approximation for that group. This lower resolution model with reduced degrees of freedom permits an increase in simulation times or system size (Figure 5.1) (Kmiecik *et al.*, 2016).

A number of different programs that simulate protein flexibility and movement, such as CABS-flex (Jamroz *et al.*, 2013) and pdb2movie (Jimenez-Roldan *et al.*, 2012; Römer *et al.*, 2016) are also available.

Proteins rarely act in isolation (Xenarios *et al.*, 2002), often locating and binding (docking) to specific binding partners in a very precise fashion, despite being in a dense and crowded environment (Vakser, 2014). Protein dimerisation and formation of higher order oligomers is very common and facilitates many biological processes such as enzyme activation or inhibition, allosteric regulation, substrate induced cooperativity of ligand binding, cell signalling and transport across cell membranes (Marianayagam *et al.*, 2004). Cytoplasmic proteins, such as aaRSs, are active in homo- and hetero- dimeric or multimeric forms (Naganuma *et al.*, 2009; Logan *et al.*, 1995; Mosyak *et al.*, 1995), other proteins such as tryptophan synthetase form heterodimers with intramolecular tunnels for substrate channelling (Hyde *et al.*, 1988; Wheeldon *et al.*, 2016). In addition integral membrane proteins FtsL, FtsB and FtsQ which are involved in cell division have been shown to form complexes in the membrane (Buddelmeijer and Beckwith, 2004; Glas *et al.*, 2015). Proteins also commonly form interactions with other macromolecular structures. In the cytoplasm, elongation or initiations factors, aaRSs and DNA-binding proteins, interact with ribosomal RNA, tRNA and DNA respectively (Ramakrishnan, 2002; Schmeing *et al.*, 2009; Wang *et al.*, 2015; Luscombe *et al.*, 2000) whilst many amphitrophic proteins form reversible interactions with the membrane (Johnson and Cornell, 1999).

Whilst, X-ray crystallography is often considered the most advanced method for obtaining high resolution structural information of proteins, it is not always possible to generate high-quality crystals (Acharya and Lloyd, 2005). Other techniques such as Nuclear Magnetic Resonance (NMR) spectroscopy and Cryo-Electron Microscopy (EM) are increasingly being used in the field of structural biology to overcome or circumvent some of the challenges faced by X-ray crystallography, however, all techniques have their limitations. In the absence of successful structure determination using experimental techniques, structural modelling and computational simulations can be useful (Krieger *et al.*, 2003). As such a number of programs and web-based servers have been developed to facilitate predictions of protein structure using methods such as homology modelling, threading and *ab*

*initio* structure prediction. This moves beyond simple sequence based prediction of secondary protein structural characteristics including trans-membrane helices and signal peptides.

## 5.2 Aims

The aim of this chapter was to use existing modelling techniques and the newly solved structure of a homologue protein to generate an improved model of MurM and identify potential Lipid binding sites. Simulations investigating protein:protein and protein:membrane interactions were also conducted and the impact of membrane phospholipids (cardiolipin (CL) and phosphatidylglycerol) on MurM were characterised, in order to gain a better understanding of the mechanism of action of this enzyme. The knowledge from these studies could be used to generate hypotheses, direct future studies and assist experimental design.

## 5.3 An improved homology model of $MurM_{159}$

### 5.3.1 Identification of an improved template for homology modelling

Crystal structures of two functional homologues of MurM have been solved previously; *S. aureus* FemA (Benson *et al.*, 2002) and *W. viridescens* FemX (Fonvielle *et al.*, 2013; Biarrotte-Sorin *et al.*, 2004). In addition, the structure of *S. aureus* FemX, which not only utilises the same Lipid II(Lys) substrate but also appends the first amino acid of the cross-bridge, has been solved recently in our lab (unpublished). Pairwise sequence alignment of *S. aureus* FemA, *S. aureus* FemX and *W. viridescens* FemX with $MurM_{159}$ indicates that the sequence identity of these homologues to MurM is low, however, *S. aureus* FemX showed the highest sequence identity to $MurM_{159}$ with 27.73 % (Table 5.1).

| **Homologue** | **Sequence Identity (%)** | **Sequence Similarity (%)** |
|---|---|---|
| *S. aureus* FemA | 21.6 | 40.3 |
| *S. aureus* FemX | 27.3 | 44.3 |
| *W. viridescens* FemX | 23.1 | 37.5 |

**Table 5.1: Sequence identity and sequence similarity between *S. pneumoniae* (159) MurM and three homologues.** *S. pneumoniae* MurM sequence was aligned with three homologues *S. aureus* FemA, *S. aureus* FemX and *W. viridescens* FemX using EMBOSS Needle (Madeira *et al.*, 2019).

The previous homology model of MurM (Fiser *et al.*, 2003) used *S. aureus* FemA, which had a sequence identity of 21.6 %. Whilst *S. aureus* FemA, similarly to *S. pneumoniae* MurM, is an aminoacyl transferase, it is functionally different in three key ways (Schneider *et al.*). Firstly, *S. aureus* FemA is responsible for the invariable addition of glycine to Lipid II(Lys) whereas MurM is responsible for addition of either L-Ala or L-Ser (Lloyd *et al.*, 2008). Secondly, FemA is responsible for the addition of two amino acids in a pentapeptide cross-bridge whilst MurM is only responsible for the addition of only one amino acid in a dipeptide cross-bridge. Finally, FemA catalyses the addition of amino acids at positions 2 and 3 of the pentapeptide cross-bridge, therefore accepting Lipid II(Lys) with a single amino acid branch as its substrate, whilst in contrast MurM can only append an amino acid at position 1 of the dipeptide cross-bridge to its substrate Lipid II(Lys) (Lloyd *et al.*, 2008).

In contrast to FemA, *S. aureus* FemX, has a higher sequence identity at 27.3 %, and is a closer functional homologue to $MurM_{159}$. Whilst *S. aureus* FemX does utilise a different tRNA-donor (glycyl-$tRNA^{Gly}$), it is similarly responsible for the addition of only one amino acid at the first position of the cross-bridge. Lipid II(Lys) is therefore the substrate for both *S. aureus* FemX and *S. pneumoniae* $MurM_{159}$. The position at which an aminoacyl transferase inserts amino acids is clearly highly specific, since FemA and FemB are not able to substitute each others activity *in vivo* despite having 37 % sequence homology and both being responsible for the

addition of two glycines to the pentapeptide cross-bridge, at positions 2 and 3, and 4 and 5 respectively (Schneider *et al.*; Ehlert *et al.*, 1997). It is likely that differences in the active sites, generate specificity for the Lipid II(Lys) substrate resulting in the specific functional characteristics of these aminoacyl-transferases.

*S. aureus* FemX and *S. pneumoniae* MurM have a sequence identity which falls into the 'safe homology modelling zone', which indicates the likelihood of two proteins sharing similar structures based upon their percentage sequence identity (Krieger *et al.*, 2003). In addition FemX has a high functional homology to MurM, therefore it was decided that *S. aureus* FemX would be used as a template for homology modelling of MurM.

### 5.3.2 $MurM_{159}$ homology model

Initial MurM modelling was conducted in collaboration with Dr. Charo I. del Genio (School of Life Sciences, University of Warwick). Figure 5.2 shows the top scoring model of $MurM_{159}$, generated using *S. aureus* FemX as a template. The predicted structure of MurM corroborates previous structural predictions (Fiser *et al.*, 2003) in that it consists of a globular domain that comprises two twisted $\beta$-sheets cores surrounded by alpha helices, and a coiled-coil helical domain.

Figure 5.3 (a) and (b) highlight some key structural differences between the existing MurM model (Fiser *et al.*, 2003) and the new MurM model. There is a loss of secondary structure at the N-terminus and increase in secondary structure at the C-terminus in the new model when compared to Fiser's existing model. Three previously unstructured regions in the globular domain of the existing model show secondary structure in the new model and the coiled-coil region ($\alpha$10 and $\alpha$11) is structured symmetrically with a small unstructured region forming a tight bend at the tip in the new MurM model. The most notable difference when comparing the two models of MurM is the location of the putative Lipid II(Lys) binding site. Figure 5.3 (c) and (d) use electrostatic surface visualisation and protein orientation

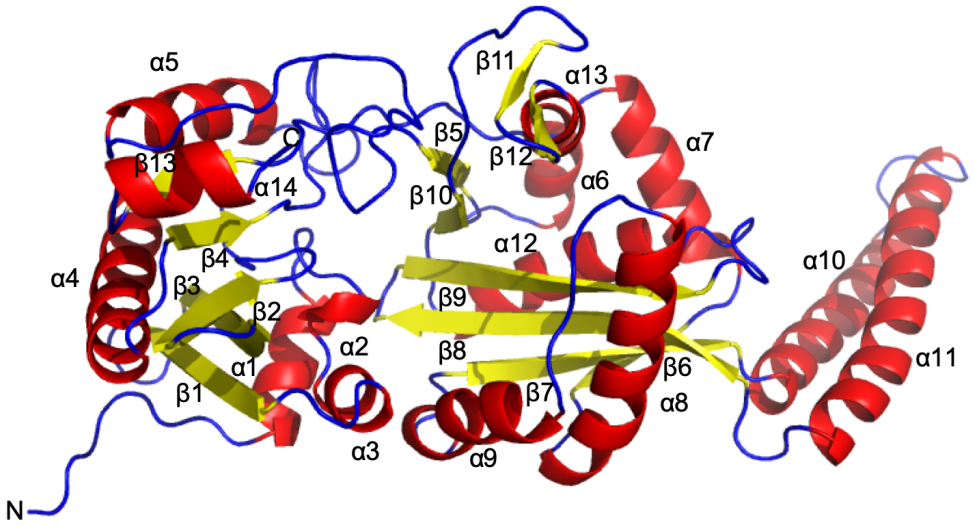



**Figure 5.2: Cartoon representation of gMurM$_{159}$ predicted structure.** 14 $\alpha$-helices (red), 12 $\beta$-sheet (yellow) and unstructured regions (blue). Best model obtained based on SOAP and DOPE scores following homology modelling using MODELLER with *S. aureus* FemX as a template.

identical to that in (Fiser *et al.*, 2003), in order to make direct comparisons between the two MurM models.

The location of the predicted peptide binding site in the Fiser *et al.* (2003) MurM model was proposed upon identification of a structural relationship between MurM domain I and the N-myristoyl transferase (NMT) fold. Four key features were used in support of identification of the predicted binding site such as; the presence of a Lipid substrate for both MurM and NMT's, the importance of the C-terminus in both MurM and NMT's, the identification of a negatively charged surface patch in MurM, and the structural location of the most highly conserved region of MurM. Each of these considerations will be challenged and it is proposed here that the Lipid II(Lys) binding site of MurM is more similar to that of *W. viridescens* FemX.

Whilst the substrate of both NMT and MurM are lipid substrates, they are contextually and structurally very different. The NMT proteins are soluble cytoplasmic proteins, which contain a deep, narrow pocket which is highly specific for the myristoyl fatty acyl chain (Wright *et al.*, 2010; Heuckeroth *et al.*, 1988).

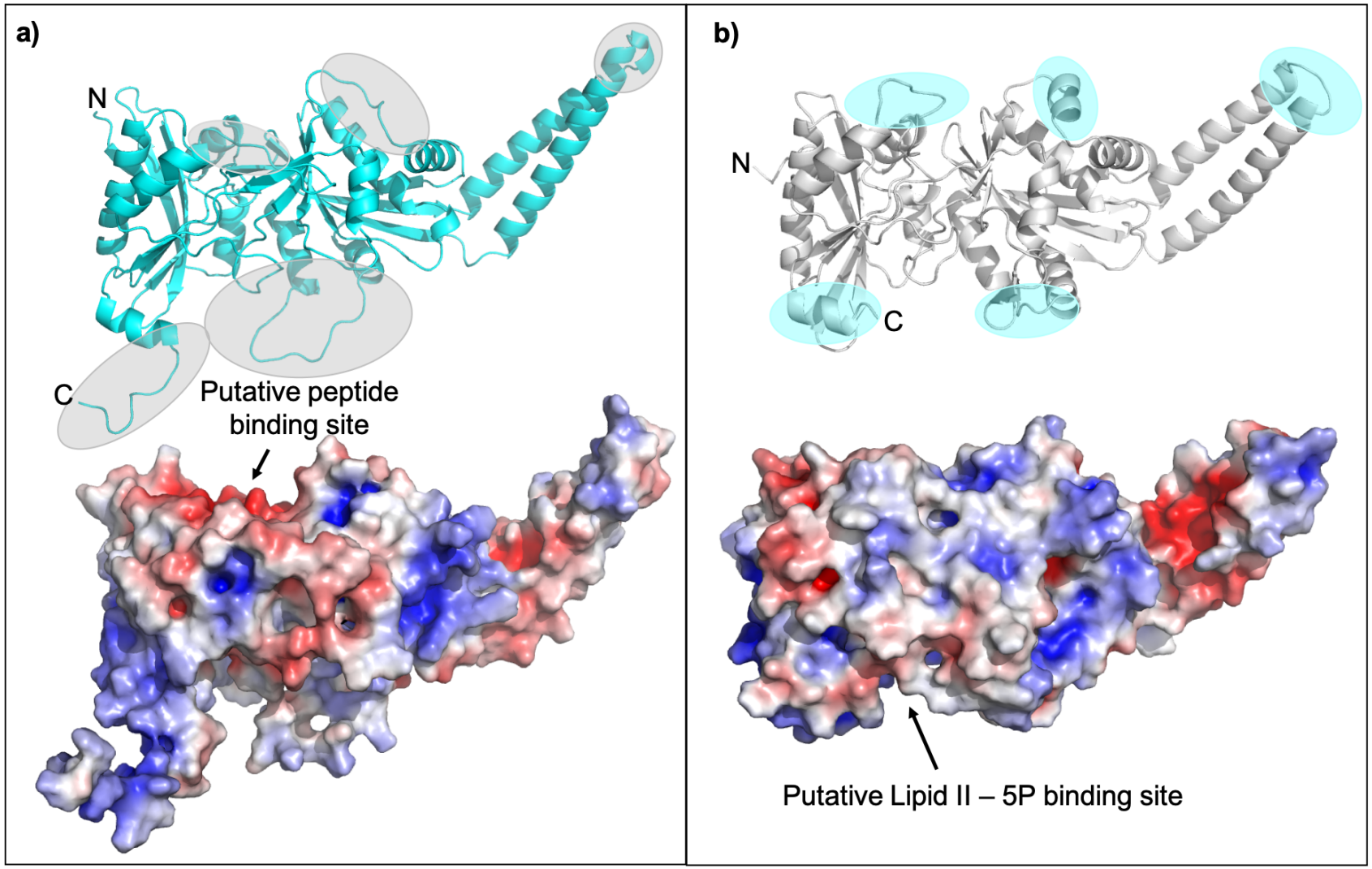



**Figure 5.3: Cartoon and Surface comparisons of previous and current MurM models.** Panel a) previous MurM model (Fiser *et al.*, 2003), b) current MurM model. Models are displayed in the orientation shown in (Fiser *et al.*, 2003) for direct comparison. Key differences are highlighted in the cartoon representation and putative binding sites indicated on the electrostatic surface representation. Figure prepared in PyMOL (Version 2.2.0) using the APBS Electrostatics Pluggin.

In contrast, since the undecaprenyl C55 lipid tail of the Lipid II(Lys) substrate is buried in the membrane, MurM does not interact with this. Instead since MurM appends L-Ser or L-Ala to the third position lysine of the pentapeptide, it is the headgroup (MurNAc-5P) which must be available for MurM binding. The headgroup is oriented such that the pentapeptide chain is in an extended conformation, perpendicular to the membrane. Since the estimated vertical dimension is only ∼1.9 nm (Ganchev *et al.*, 2006), the MurM protein must be in very close proximity to the lipid bilayer. Given the stark contextual difference of the Lipid substrate for MurM, it seems unlikely that this protein would have a binding site similar to that of an NMT.

The C-terminus in NMT proteins is important for catalysis (Farazi *et al.*, 2001) whilst the C-terminus has been shown to be required for MurM activity (Filipe *et al.*, 2001b), it remains unknown as to whether this is due to structural or functional reasons. The C-terminus of the Fiser *et al.* (2003) MurM model was unstructured due to absence of electron density in the template, however, in the new MurM model the C-terminus is a short $\alpha$-helix which ends at the edge of the proposed binding site. The C-terminus is therefore in the immediate vicinity of the putative Lipid II(Lys) binding site and as such could play a functional or structural role.

Finally, Fiser *et al.* (2003) identified 11 residues, between residue 66 and 84, which were fully conserved between FemA and MurM. Since these 11 residues also overlapped with the negatively charged surface patch, they were believed to form a dominant part of the proposed binding site (Fiser *et al.*, 2003). These 11 residues are not well conserved between *S. aureus* FemX and the new model of MurM, but are relatively highly conserved between different strains of *S. pneumoniae* MurM. Since the negatively charged surface is still present in the new MurM model, and is located on the opposite side of the protein to the newly proposed Lipid II(Lys) binding site, it may instead be important for protein:protein interactions.

Further evidence and characterisation of the newly proposed Lipid II(Lys) binding site is outlined in this chapter demonstrating that the new MurM model provides a

more accurate and informative representation of MurM then the previous model.

Figure 5.4 shows a Ramachandran plot of the new MurM model indicating that 91.3 % of all residues are within favoured regions and 97.0 % of all residues are in allowed regions. The 12 outliers are Leu64, Gly65, Glu123, Ile133, Val139, Asp263, Thr265, Arg266, Asn297, Asp327, Asn369 and Tyr387. Whilst this model has not achieved the desired $\leq 0.05\%$ of outliers and $\geq 98\%$ of residues in favoured regions, a small amount of uncertainty is expected due to the nature of these techniques and it has been noted that even well established crystal structures do not always achieve these desired values.

### 5.3.3 Identifying a possible Lipid II(Lys) binding site

Biarrotte-Sorin *et al.* (2004) solved the structure of *W. viridescens* FemX with the co-crystallised UDP-MurNAc-5P substrate bound which identified 9 residues making contact with the substrate. Similarly Fonvielle *et al.* (2013) co-crystallised and solved the structure of *W. viridescens* FemX in the presence of a peptidyl-RNA conjugate and identified 10 residues involved in binding the peptide substrate. Figure 5.5 shows the differences between the two substrates used. When the *W. viridescens* FemX crystals are aligned the substrates overlay well, indicating that the two structures corroborate the others findings. Since the UDP-MurNAc-5P is closer structurally to the Lipid II(Lys) substrate of MurM, this was used initially to investigate the MurM binding site.

*W. viridescens* Fem X is 340 amino acid residues long compared with MurM which is 403 residues, it lacks the coiled-coil helical domain, and has low sequence homology with MurM. However despite this, MurM does, with very low catalytic activity, utilise UDP-MurNAc-5P, the substrate of *W. viridescens* FemX. Since the root mean squared deviation (RMSD) of *W. viridescens* FemX and the new MurM model was 2.355 Å, the structure of *W. viridescens* FemX with UDP-MurNAc-5P was used to assist identification of the MurM Lipid II(Lys) binding site.

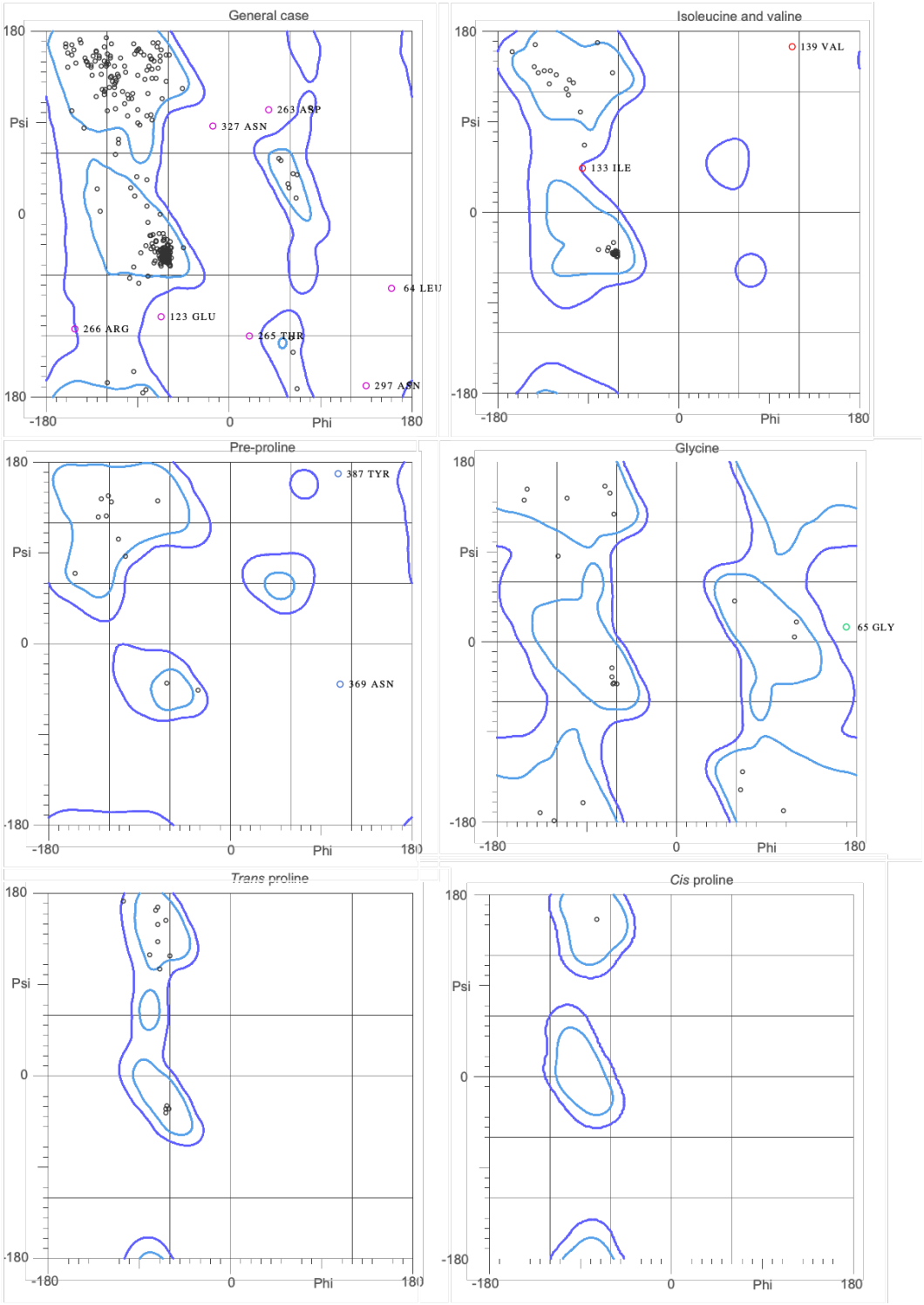



**Figure 5.4: Ramachandran Plot of MurM159 model.** Shows 91.3 % of all residues are within favoured regions (light blue lines) and 97.0 % of all residues are in allowed regions (dark blue lines). There are 12 residues that are outliers (pink circles). Plot generated in MolProbity (Lovell *et al.*, 2003).

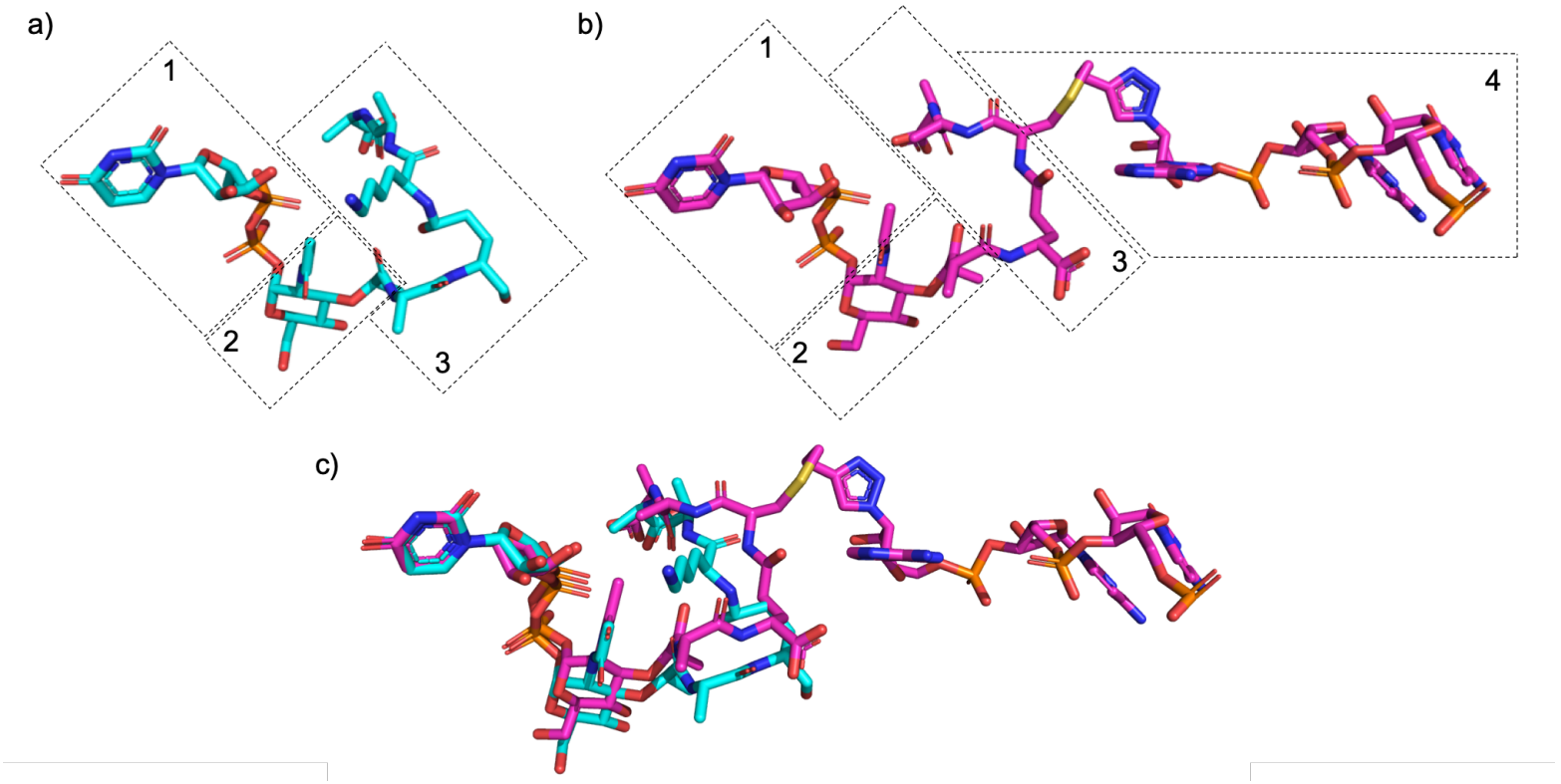



**Figure 5.5: Comparison of UDP-MurNAc-5P substrates in the *W. viridescens* FemX binding site.** a) stick structure of UDP-MurNAc-5P as found in FemX (Biarrotte-Sorin *et al.*, 2004), b) stick structure of UDP-MurNAc- peptidyl-RNA conjugate as found in FemX (Fonvielle *et al.*, 2013), c) overlay of UDP-MurNAc-5P and UDP-MurNAc-peptidyl-RNA conjugate. Box 1 = UDP, Box 2 = MurNAc, Box 3 = 5P and Box 4 = peptidyl-RNA conjugate. Figure made in PyMOL (Version 2.2.0).

Figure 5.6 shows the model of MurM which has been aligned and overlaid to that of *W. viridescens* FemX with the UDP-MurNAc-5P substrate bound. It can be seen that the UDP-MurNAc-5P fits well into a large cavity in the new MurM model. Since the UDP-MurNAc-5P is significantly smaller than the Lipid II(Lys) substrate of MurM (Appendix 7.5) this can only be used as a guide. However, with the pentapeptide of Lipid II(Lys) bound in a similar orientation to that of the pentapeptide in the structure of *W. viridescens* FemX, the prenyl chain of Lipid I(Lys) would protrude out of the active site and away from the protein into the lipid bilayer.

The following 8 residues; Tyr103, Lys36, Asn38, Trp39, Thr209, Arg211, Tyr215, Tyr256, were independently identified in both *W. viridescens* FemX structures to be involved in substrate binding (Biarrotte-Sorin *et al.*, 2004; Fonvielle *et al.*, 2013). Amino acid residues with similar properties which are closely related in sequence location or physical space and orientation were identified as 'corresponding residues'. The corresponding residues in $MurM_{159}$ were therefore identified in MurM as; Phe103, Lys35, Asn/Asp37, Trp38, Thr206, Arg215 and Tyr219. Pairwise alignment $MurM_{159}$ from 9 different strains (Figure 5.7) shows the sequence location and conservation of the equivalent residues in *S. pneumoniae*.

All residues, except Asn/Asp37 were conserved across all strains investigated. In both penicillin sensitive strains (Pn16 and R6) the residue was Asn37, whilst in all resistant strains a Asp37 was present. Residues Phe103, Lys35, Trp38, Arg215 and Tyr316 in $MurM_{159}$ were all found to occupy the same location in space as the corresponding residues in *W. viridescens* FemX, and are therefore considered structurally conserved. Overlays of the $MurM_{159}$ model and bound UDP-MurNAc-5P from the *W. viridescens* FemX crystal structure, shows that these structurally conserved residues are all pointing directly towards the substrate (Figure 5.8). It is therefore likely that these residues are involved in substrate binding in MurM similarly to those identified in Wv FemX. Additionally, whilst position 37 is not conserved between homologues from different species, in all *S. pneumoniae* strains investigated position 37 is Asn/Asp and interestingly all resistant

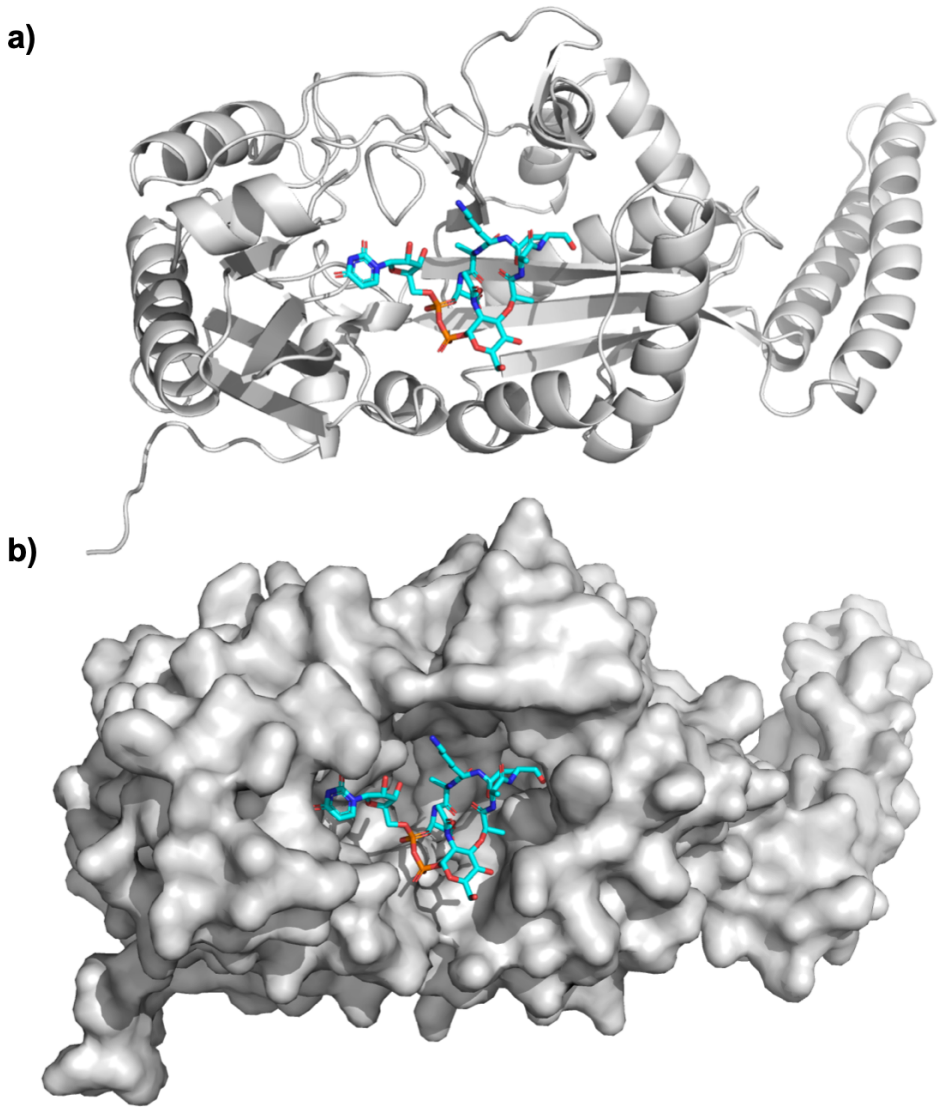



**Figure 5.6: Current MurM model with UDP-MurNAc-5P overlaid.** a) Cartoon representation of MurM with UDP-MurNAc-5P overlaid (cyan) b) surface representation of MurM with UDP-MurNAc-5P overlaid. UDP-MurNAc-5P can be seen to fit tightly into the cavity shown, this is the proposed binding site for the Lipid II(Lys) substrate of MurM. Figure prepared in PyMOL (Version 2.2.0).

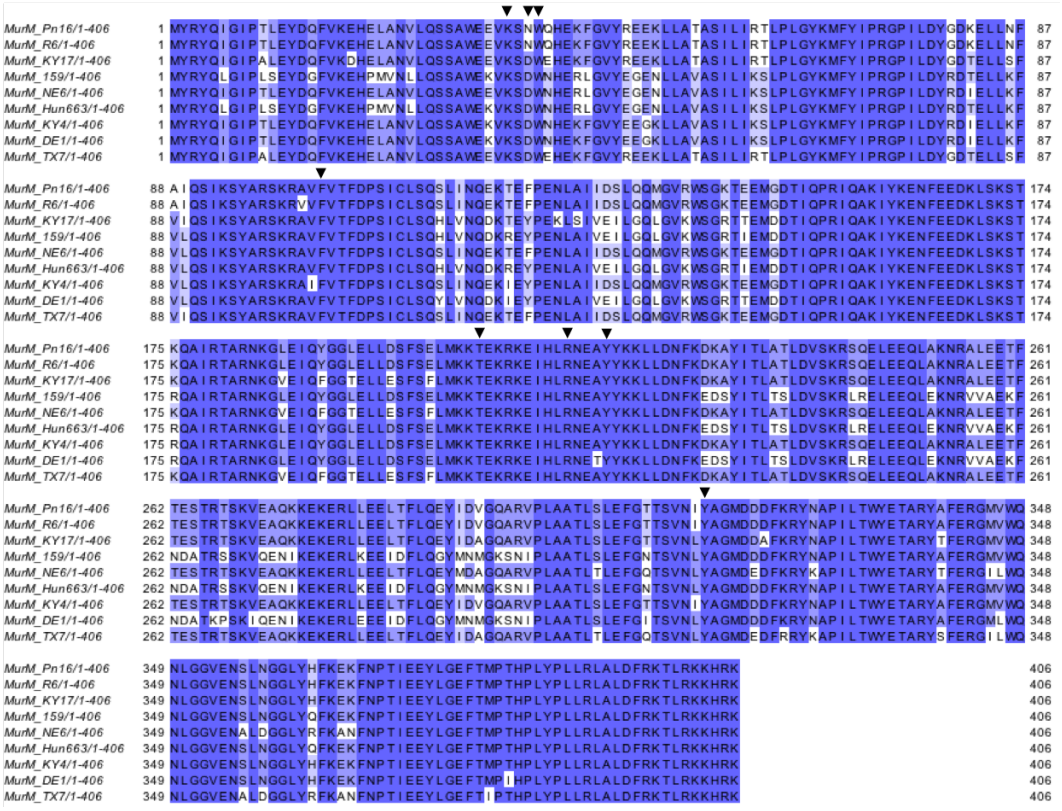



**Figure 5.7: Pairwise sequence alignment of MurM sequences.** Black arrows indicate residues proposed to be involved in substrate binding. Residues coloured according to % identity using Jalview (Version 2.10.5).

strains were found to contain Asp at this position. Thr206 was not found to be facing the substrate binding cavity and was at right angles with the equivalent Thr209 in *W. viridescens* FemX and is therefore unlikely to be important for substrate binding in MurM.

An additional three residues Trp31, Asp149 and Glu366 were identified as having potential importance in the MurM binding site. All three residues point directly towards the cavity where the substrate binding is proposed to occur, Trp31 is fully conserved among all *S. pneumoniae* strains investigated whilst Asp149 and Glu366 are only partially conserved.

Asp109 was identified by site directed mutagenesis as a possible catalytic residue in *W. viridescens* (Hegde and Blanchard, 2003), in MurM the Asp107 is similarly conserved (Lloyd *et al.*, 2008). Asp107 is structurally conserved, Figure 5.8 shows it oriented directly towards the substrate, with a distance of 12.5 Å between Asp107 and the epsilon amino group of the third position lysine. The surface structure of the new MurM model indicates that interaction of Asp107 with lysine may be partially obstructed by some unstructured regions of the enzyme. It is therefore possible that during catalysis structural rearrangement may increase the proximity of the catalytic Asp107 relative to the Lipid II(Lys) substrate. The *W. viridescens* FemX with bound UDP-MurNac-5P has been overlaid with MurM in order to assist with identification of the substrate binding site and important residues. However, the UDP-MurNac-5P is significantly smaller than Lipid II(Lys) substrate (Appendix 7.5) of MurM and can therefore only be used as a guide. With Lipid II(Lys) bound in a similar orientation to the UDP-MurNAc-5P in Figure 5.8, the Lipid chain of Lipid II(Lys) would protrude out of the active site.

Residues 244-274 of the coiled-coil region have been previously identified as important for determining the specificity of the amino acid to be incorporated into the PG (Filipe *et al.*, 2001b). Additionally, the role of the coiled-coil arm for tRNA binding in *T. thermophilus* SerRS has been demonstrated (Biou *et al.*, 1994). It is therefore likely that the coiled-coil arm of MurM also plays a role in tRNA

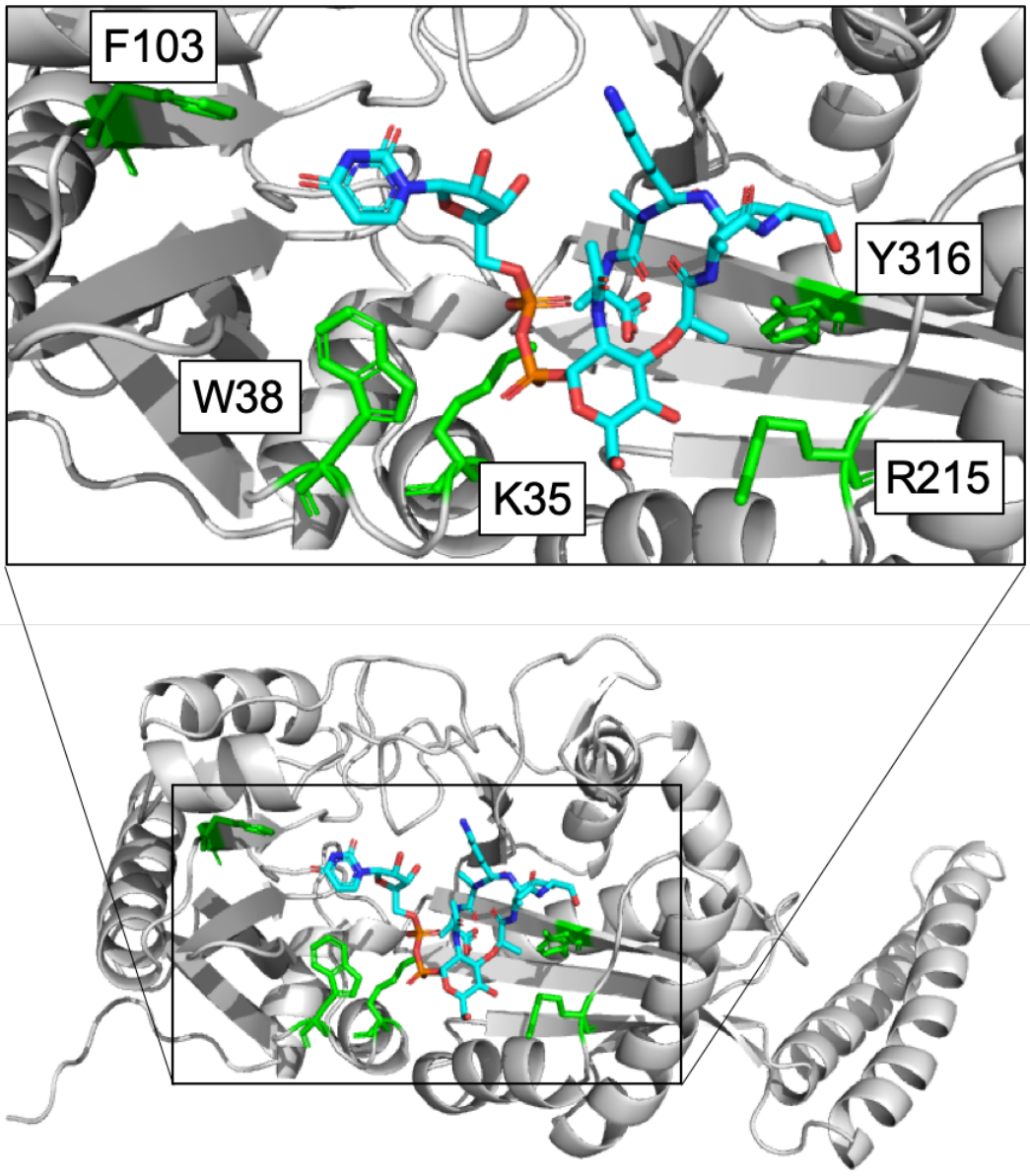



**Figure 5.8: Potential $MurM_{159}$ binding site residues.** MurM with UDP-MurNac-5P overlaid, inset close up of proposed binding site with residues F103, K35, W38, R215 and Y316 in green shown to be pointing towards the substrate. Note: single letter code for amino acids used in this diagram.

binding. Electrostatic surface visualisation of the new MurM model shows that the residues surrounding the binding site are primarily positively charged (Figure 5.9), which indicates that this area of the proteins surface may form direct interactions with the negatively charged phosphate groups of membrane phospholipids. The electrostatic surface visualisation also indicated that the negatively charged surface patch is located on the opposite side of the protein from the proposed Lipid II(Lys) binding site. Therefore, MurM may form simultaneous interactions with both the membrane and another protein.

Figure 5.10 shows regions of MurM which were previously identified to be required for activity; the 50 N-terminal residues (red), 10 C-terminal residues (blue) or residues 244-274 (green) (Filipe *et al.*, 2001a,b). Deletion of N- or C-terminal residues of MurM results in unbranched peptidoglycan which indicated that these residues are important for the branching phenotype. However it is currently unclear as to whether these residues play a functional role in enzyme activity, are important for structural integrity or correct protein folding or are important for membrane:protein or protein:protein interactions. Figure 5.10 illustrates that these regions of MurM are in close proximity to the proposed Lipid II(Lys) binding site and therefore may have a functional role. Residues 244-274 were identified as important for defining the specificity for incorporation of alanine or serine into the cross-bridge by MurM. Attempts to identify a single residue which determined specificity were only partially successful, it seems that the presence of a threonine or lysine at position 260 correlates with the incorporation of primarily serine or alanine respectively. However, the presence of that amino acid residue alone cannot determine specificity, as such additional residues are required working in tandem (Filipe *et al.*, 2001b). The importance of the coiled-coil region in determining specificity of aminoacyl-tRNA is reflected in *W. viridescens* FemX, which, has no coiled-coil domain and demonstrates a lower selectivity for aminoacylated- tRNA by its ability to utilise serine, alanine or even glycine depending upon availability. Given the importance of the coiled-coil domain in determining specificity, it is likely that this region makes contact with

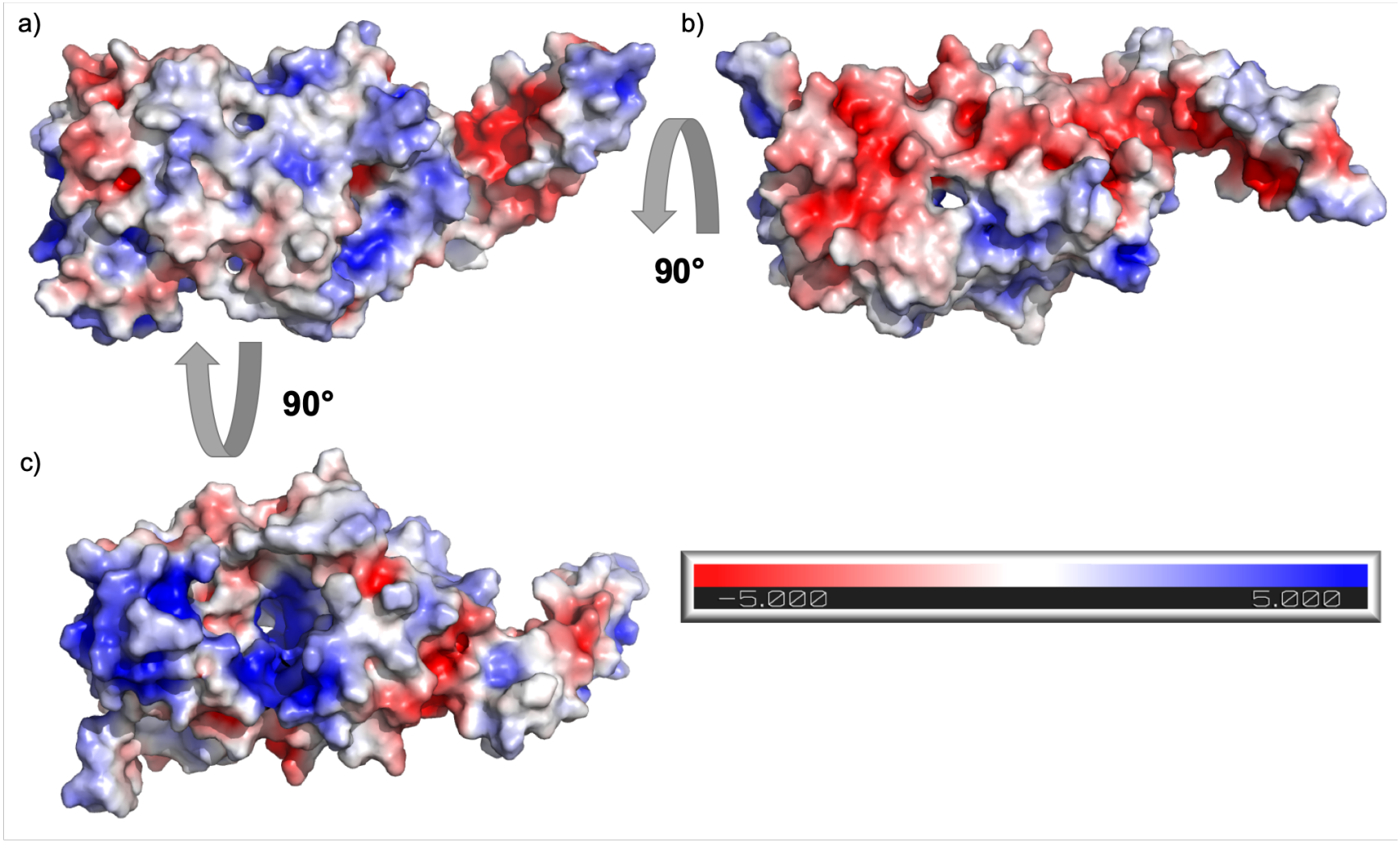



**Figure 5.9: Electrostatic surface representation of MurM$_{159}$.** a) MurM$_{159}$ b) proposed MurM$_{159}$ Lipid II(Lys) binding site c) MurM$_{159}$ negatively charged surface patch. Figure prepared in PyMOL (Version 2.2.0) using the APBS Electrostatics Pluggin.

the tRNA. MurM has previously been compared to SerRS, for which deletion of the N-terminal coiled-coil domain does not affect amino-acid activation but does reduce aminoacylation activity (Borel *et al.*, 1994), therefore indicating that it is important for interaction with tRNA. Furthermore, crystal structures of *T. thermophilus* SerRS show that the coiled-coil domain makes contact with both the variable arm and T$\psi$C loop of the seryl-tRNA (Biou *et al.*, 1994). It is possible therefore, that the coiled-coil region of MurM is similarly necessary for the formation of correct interactions with the tRNA, and plays an important part in determining the specificity of the aminoacyl-tRNA substrate.

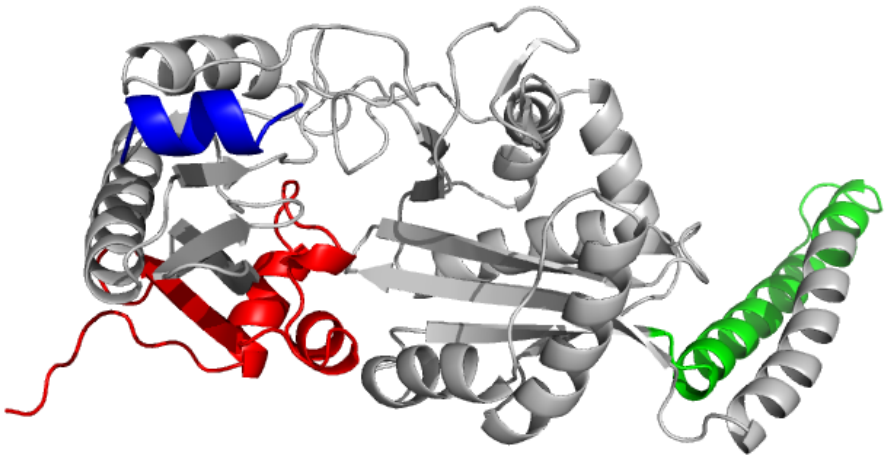


**Figure 5.10: Cartoon representation of MurM$_{159}$ with regions required for activity.** N-terminal residues (red), C-terminal residues (blue) and residues 244-274 (green) are required for MurM activity.

## 5.4 Docking of a truncated Lipid II(Lys) to MurM

Following the identification of a potential substrate binding site in the MurM model, molecular docking was used to investigate docking of the Lipid II(Lys) substrate to MurM. However, Lipid II(Lys) is a large molecule and is therefore unsuitable for molecular docking, in addition the lipid tail would be embedded in the membrane

and so is not itself part of the MurM substrate. Therefore, truncated Lipid II(Lys) with a reduced hydrocarbon tail was used for these docking experiments (Appendix 7.13).

Molecular docking of the substrate to MurM was conducted using AutoDock Vina (Trott and Olson, 2010). In the absence of defining the binding site, providing the entire protein for potential binding, the substrate docked only at the proposed binding site. Figure 5.11 shows the conformation of the substrate which minimised the free energy of the system.

Figure 5.12 shows residues which were identified to form hydrogen bonds between MurM and the Lipid II(Lys) substrate. Across all MurM sequences studied; Ser36, Lys35, and Tyr375 are completely conserved and Asp37 and Arg144 exist as conservative mutations.

Whilst the same pocket of MurM was identified as the binding site, the orientation of the Lipid II(Lys) substrate was significantly different when compared to the UDP-MurNAc-5P (overlaid in proposed MurM binding site) (Figure 5.13). Initial identification of this binding site used the *W. viridescens* FemX substrate, which, whilst similar, has a number of differences between MurMs natural substrate. In comparison to UDP-MurNAc-5P, in Lipid II(Lys) the uridine has been replaced by a long hydrocarbon chain, which is embedded in the membrane, and there is an additional GlcNAc group appended to the MurNAc. Therefore it is likely that binding of the two substrates would be slightly different. In the Lipid II(Lys) docking, the third position lysine, to which the amino acid will be transferred, is located near the opening whilst the phosphates to which the lipid tails would be appended are deeper in the binding pocket.

The diphosphates which connect to the hydrocarbon tail of the Lipid II(Lys) are located in a channel which is at a 90 ° angle to the binding site. *In vivo* the hydrocarbon of Lipid II(Lys) would be embedded in the membrane and the diphosphates would be located at the surface of the membrane. The diphosphates

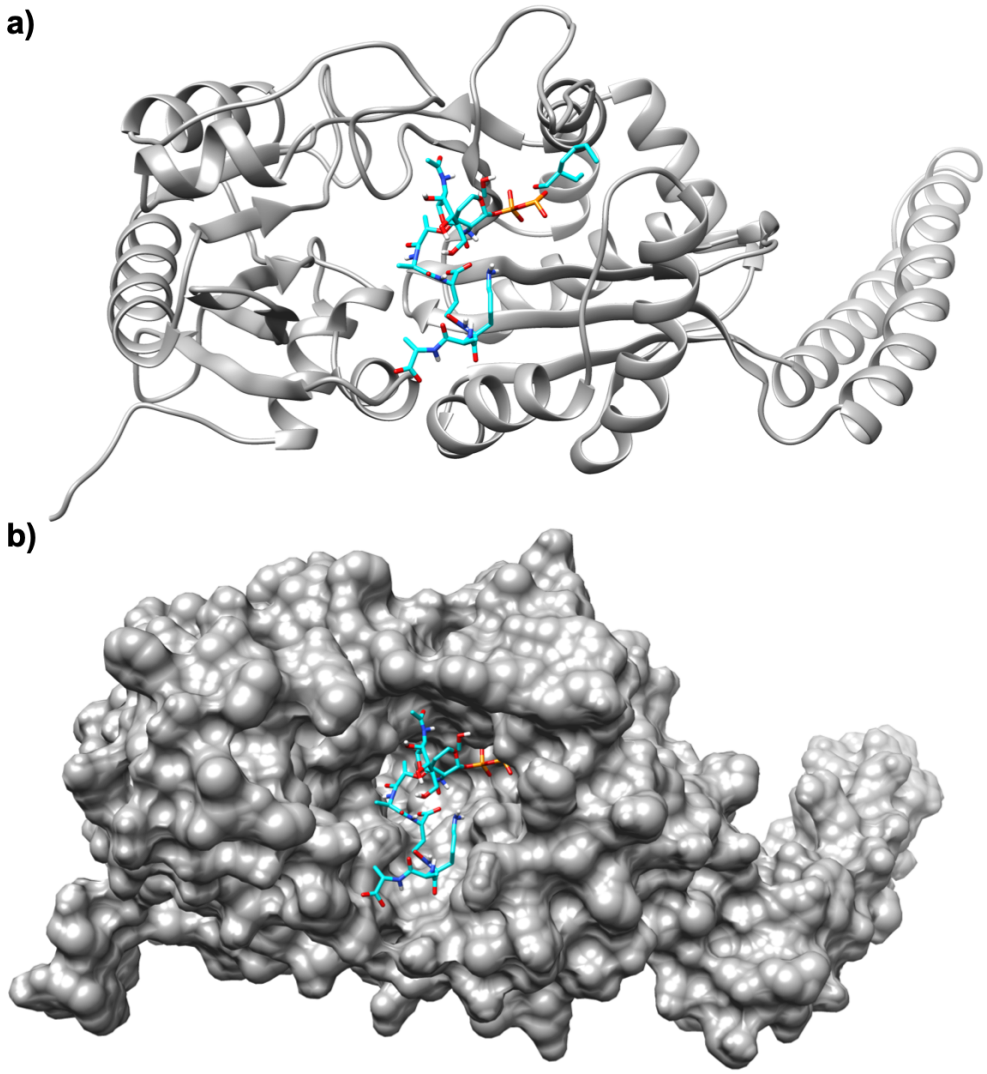



**Figure 5.11: Molecular docking prediction of current MurM model with modified Lipid II(Lys).** a) Cartoon representation of MurM with docked modified Lipid II(Lys) (cyan) b) Surface representation of MurM with docked modified Lipid II(Lys).

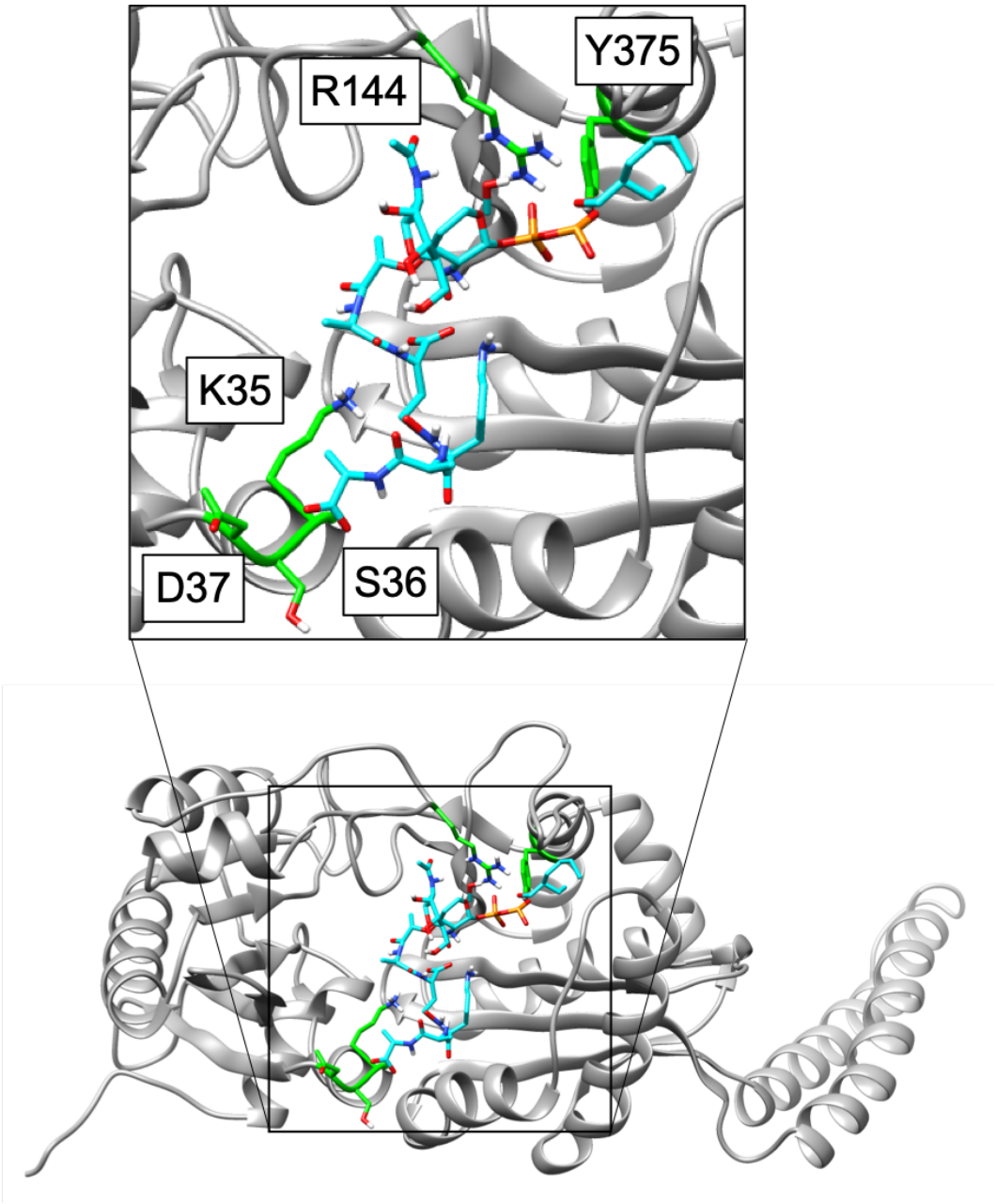



**Figure 5.12: Molecular docking predictions of MurM binding site residues.** Current MurM model with docked Lipid II(Lys). Inset: close up of residues predicted to form hydrogen bonds. Lys35, Ser36, Asp37, Arg144 and Tyr357 (green) are directed towards the substrate. Note: single letter code for amino acids used in this diagram.

in the docked substrate are ~1.5 nm deep in the binding pocket and ~2.2 nm away from the exit of the channel, therefore it is unlikely that this orientation is correct for the membrane bound Lipid II(Lys). Docking the Lipid II(Lys) in the complete absence of both the lipid tail and the diphosphates may produce results that are more plausible.

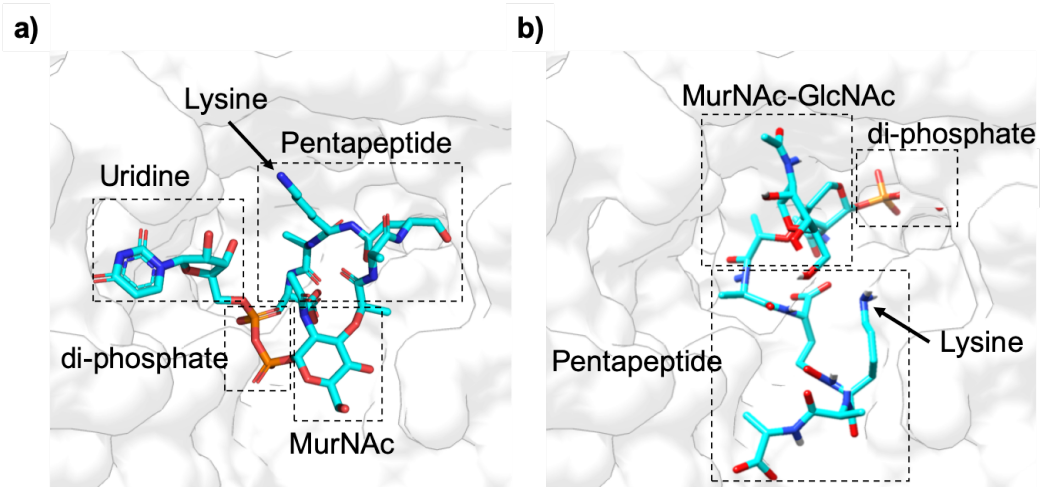



**Figure 5.13: Orientation of substrate in the binding site of current MurM model**. a) Surface representation of MurM overlaid with UDP-MurNAc-5P, b) Surface representation of MurM with molecular docking of truncated Lipid II(Lys). Dashed boxes are used to demonstrate differences between the two substrates and the relative orientations within the binding site.

## 5.5 Atomistic molecular dynamics simulations of membrane interactions with $MurM_{159}$

This work was conducted in collaboration with Prof. Syma Khalid and PhD student Jonothan Shearer (University of Southampton).

MurM is an amphitrophic protein, forming temporary and reversible interactions with the lipid bilayer, specifically through binding of its substrate Lipid II(Lys). MD techniques were used to investigate the orientation of MurM docking with the lipid bilayer and to further identify the Lipid II(Lys) binding site and any direct interactions between MurM and the membrane associated Lipid II(Lys) (Witzke *et al.*, 2016). In addition simulations were conducted with membranes containing

cardiolipin in order to identify the effect of this membrane phospholipid on MurM.

### 5.5.1 Interactions between  and the lipid bilayer

Coarse grain simulations were conducted six times for each of three membrane systems; system 1, 2 and 3 contained 75:25:0, 72:16:12 and 72:12:16 of PE, PhG and CL respectively. Interaction with the membrane was observed in $< 3$ $\mu$s and remained unchanged for the remaining 2 $\mu$s. Interaction between MurM and the membrane occurred in a number of different orientations and the predicted binding site was found both facing towards and away from the membrane. Additionally, the Lipid II(Lys) was found inside the proposed binding site in one of the simulations (Figure 5.14). Since this is a dynamic system, and these simulations cover very short periods of time, it would not be possible for all interaction scenarios to be identified. Therefore identifying Lipid II(Lys) binding in the proposed binding site in one of the 18 simulations is significant and strongly supports this as the location of the Lipid II(Lys) binding site.

Following backmapping, prevalent hydrogen bonds were identified to be occurring at Lys35, Lys60 and Arg144. Lys35 and Lys60 are completely conserved, and Arg144 exists only as conservative mutants across all MurM sequences studied. Additionally, Lys35 was also proposed to be involved in substrate binding of *W. viridescens* FemX (Biarrotte-Sorin *et al.*, 2004). Previous studies suggest that the height of the Lipid II(Lys) head group is 1.9 nm (Ganchev *et al.*, 2006); the measurement from the phosphate group at the membrane surface to the terminal carboxyl group of the D-Ala-D-Ala (at the deepest point of the binding pocket) was $\sim$1.2 nm. In addition, comparison of the orientation of UDP-MurNAc-5P substrate (overlaid in the MurM binding pocket) and Lipid II(Lys) substrate in the MurM binding pocket, shows that the positioning of the diphosphate, MurNAc and pentapeptide (and particularly the third position lysine) is strikingly similar (Figure 5.15). Taken together this strongly supports the identification of the binding pocket and suggests that the Lipid II(Lys) binds to MurM in this orientation which is similar to that of *W.*

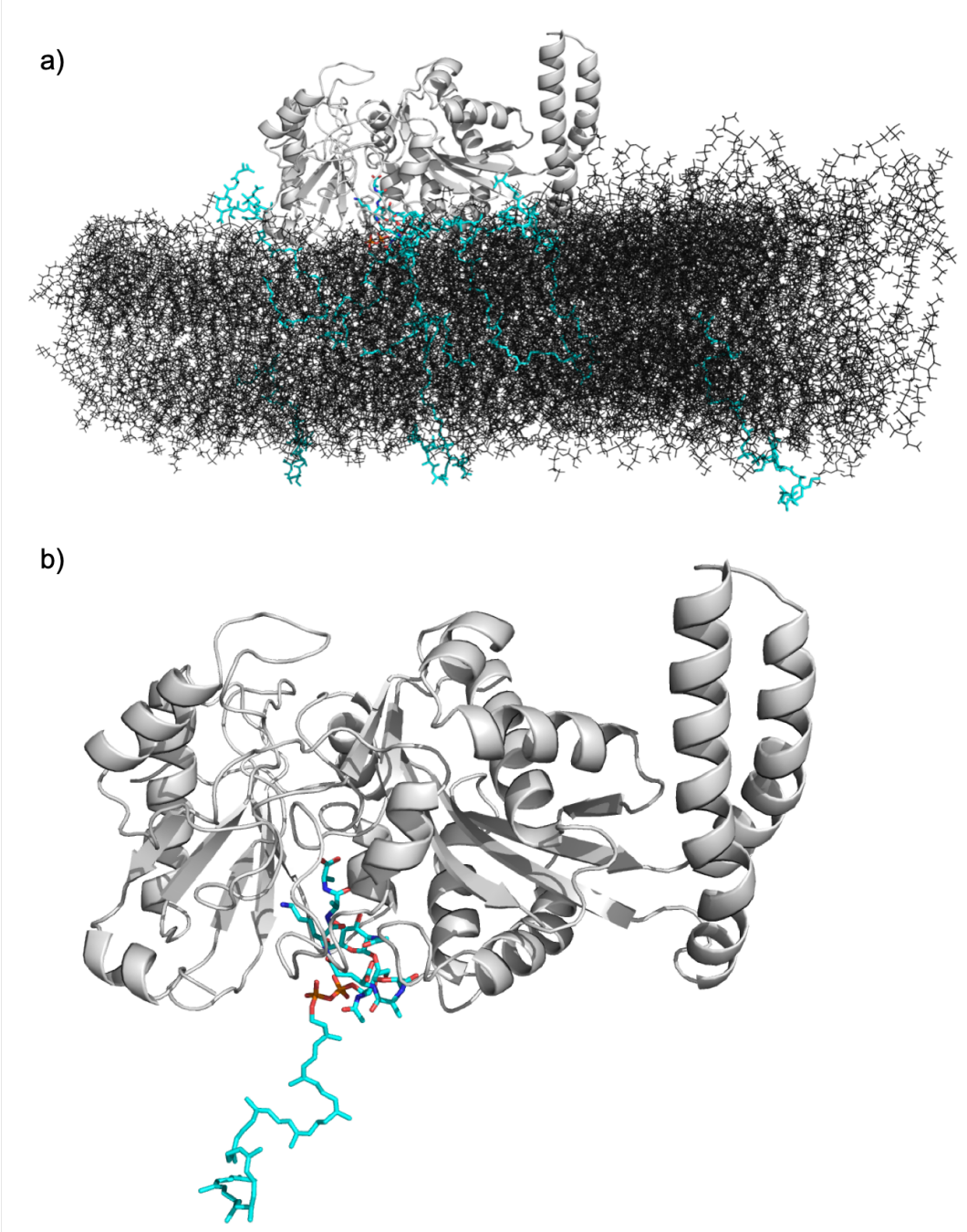



**Figure 5.14: Interactions between current MurM model and the cell membrane.** a) MurM association with the membrane containing Lipid II(Lys) (cyan), b) Lipid II(Lys) (cyan) in the binding site of current MurM model (cartoon representation).

*viridescens* FemX binding to its substrate.

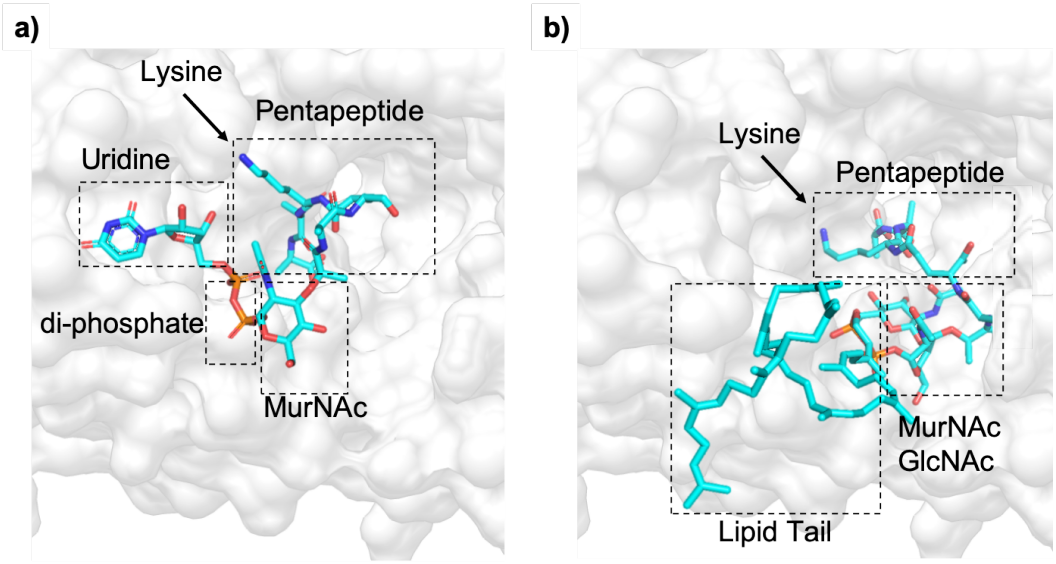



**Figure 5.15: Orientation of Lipid II(Lys) in the binding site of current MurM model**. a) Surface representation of MurM overlaid with UDP-MurNAc-5P, b) Surface representation of MurM with membrane embedded Lipid II(Lys). The surrounding membrane has been removed for clarity. Dashed boxes are used to demonstrate differences between the two substrates and the relative orientations within the binding site.

In each membrane system, two repeats, one with the binding site facing toward the membrane and one with the binding site facing away from the membrane, were backmapped to the all-atom model (Appendix 7.10). All-atom simulations indicated that in general, Lipid II(Lys) clustered around the MurM primarily associating with the positively charged surfaces.

### 5.5.2 Interactions of $MurM_{159}$ with phospholipids

Phospholipids such as CL and PhG are known to be involved in the spatio and temporal biochemistry of cells. The presence of cardiolipin has been demonstrated to upregulate the activity of MurG (Lin *et al.*, 2019; Boots *et al.*, 2003) and similarly the activity of $MurM_{159}$ *in vitro* was reported to increase in the presence of CL and decrease in the presence of PhG (unpublished data - Dr Adrian Lloyd). We therefore investigated the effects of CL and PhG to the MurM simulations. A Depletion-Enrichment (D-E index) was used to measure enrichment or depletion of

specific lipids in the membrane occurring within a 1.1 nm perimeter of the protein. In the absence of CL (membrane system 1) the D-E index of PE and PhG is is $\sim$1, indicating that upon MurM binding to the membrane, there is no change in distribution of PE or PhG throughout the membrane (Appendix 7.9). In the presence of CL, the D-E index for PE and PhG remains similar or decreases slightly (for PhG); concomitantly, an increase the D-E index of CL was observed, indicating that CL moves into the area of membrane making contact with MurM. Figure 5.16 shows example data of the D-E index and a visual representation of CL enrichment around MurM, from both membrane systems. CL enrichment occurs in simulations with different MurM orientations, and also in both membrane systems; CL enrichment is slightly more pronounced in membrane system 2 (with 12 % CL). These experiments demonstrate a relationship between phospholipids and MurM; the increase in activity of MurM in response to CL, and the enrichment of CL around MurM at the membrane go hand in hand. Given that MurM successfully associates with a membrane void of CL, it is possible that CL may be recruited to the membrane location of MurM following its association, in order to regulate its activity. Further investigation into the specific sites of MurM which interact with CL would be beneficial.

## 5.6 $MurM_{159}$ Flexibility Modelling

Flexibility modelling was conducted in collaboration with Prof. Rudolf Roemer (University of Warwick) utilising a fast and efficient method for generating protein flexibility simulations which could provide insight into functionality (Jimenez-Roldan *et al.*, 2012; Römer *et al.*, 2016).

Determining the flexibility of MurM showed that the coiled-coil helical domain is highly flexible. Modes 7 and 8 show flexibility of the coiled-coil in the z and x plane whilst mode 10 shows rotation about the unstructured region that connects it to the globular domain. Conformational adaptation of the coiled-coil helical arm

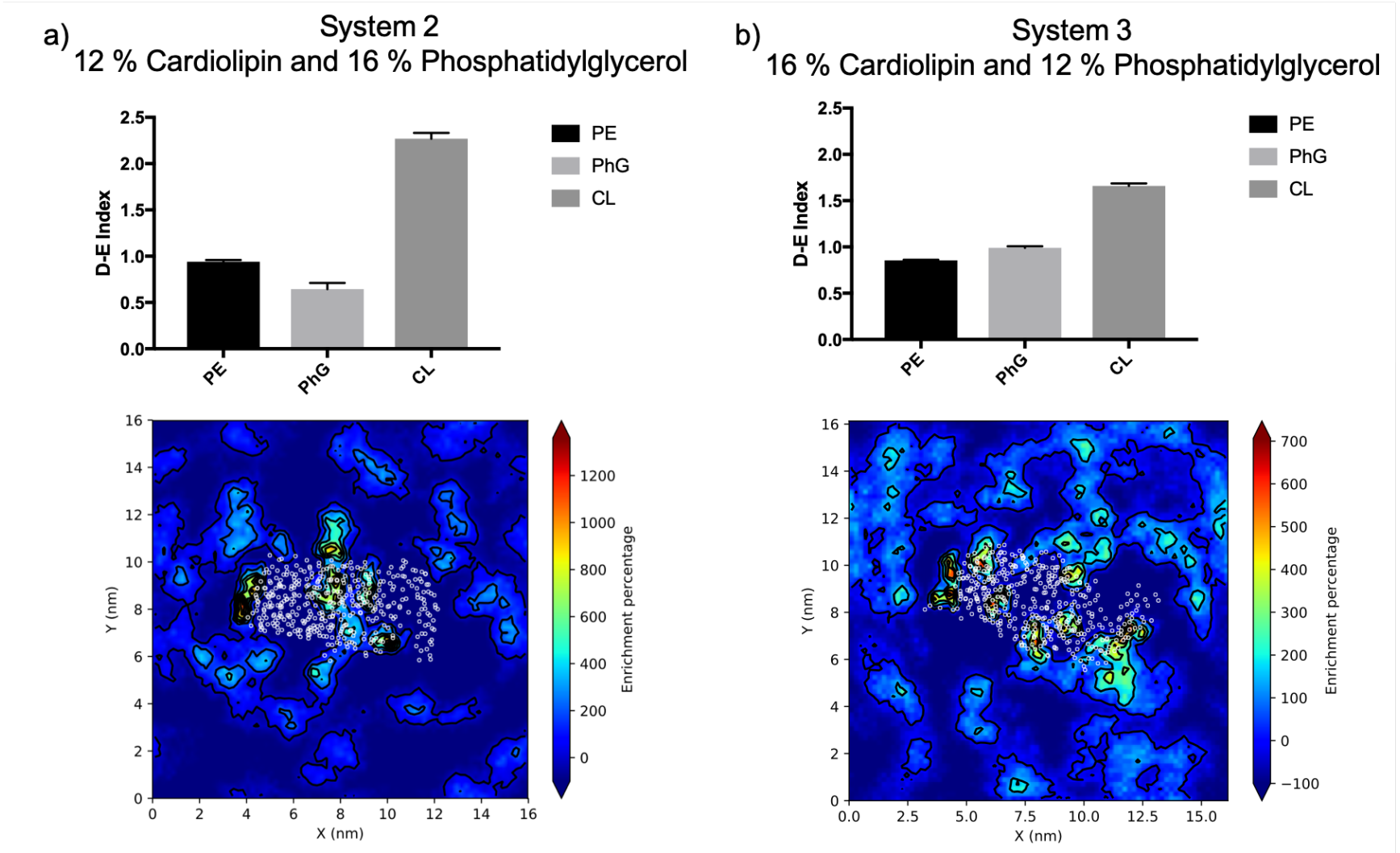



**Figure 5.16: Depletion-Enrichment Index of phosphatidylethanolamine, phosph- atidylglycerol and cardiolipin with MurM:membrane associate.** a) Membrane system 2 and b) Membrane system 3. TOP: Histogram demonstrating the relative D-E index of phosphatidylethanolamine (PE), phosphatidylglyerol (PhG) and cardiolipin (CL) within 1.1 nm of MurM. Enrichment percentage >1 indicated that lipid is enriched, whilst <1 indicates that lipid is depleted compared to average lipid density. GraphPad Prism (Version 7.0c) was used for data analysis and figure preparation. BOTTOM: Enrichment map of cardiolipin around MurM at the membrane surface. White dots represent the center of geometry of each protein residue. All simulations with MurM in different starting orientations generated similar data (Appendix 7.9). Figures prepared using Matplotlib (Version 3.0.3).

has been suggested to be of importance in the recognition of tRNA for human SerRS (Biou *et al.*, 1994). Given the structural similarities between MurM and SerRS in combination with the known importance of residues 224-274 in MurM for determining amino acid specificity, it is possible that the coiled-coil helical arm of MurM also undergoes conformational changes to facilitate tRNA binding. However, unlike the SerRS which requires a long variable arm of tRNA to make interactions with the coiled-coil arm (Biou *et al.*, 1994), it has previously been shown that MurM can utilise a mini-helix (Lloyd *et al.*, 2008), containing only a acceptor stem and a T$\psi$C stem, meaning that exact formation of tRNA interactions differ between MurM and SerRS.

Mode 10 identified movements resembling the closing of the binding site in the globular domain. This may for example, represent a conformational change that occurs during binding of Lipid II(Lys), or a conformational change that occurs during catalysis.

## 5.7 Prediction of protein:protein interactions

Previously bacterial two hybrid assays showed that *S. aureus* FemA and FemB both form homo- and hetero-dimers, *S. pneumoniae* MurN, but not MurM, was shown to form homodimers and MurM and MurN did not form heterodimers. This experiment used MurM and MurN as 'prey' and 'bait' proteins, each of which was conjugated to a domain of the *Bordetella pertussis* CyaA protein. As such, it is possible that the folding of MurM and MurN was altered by the additional CyaA domain resulting in aberrantly folded protein, or that the presence of the CyaA domain itself hindered any possible homo- or hetero-dimerisation of these proteins. In addition, unlike the positive results, the negative results were not corroborated using additional methods (Rohrer and Berger-Bchi, 2003). Since interactions were found to occur between FemA and FemB, and given the observation of a highly negative patch on the surface of MurM, ZDOCK server (Pierce *et al.*, 2014) was used

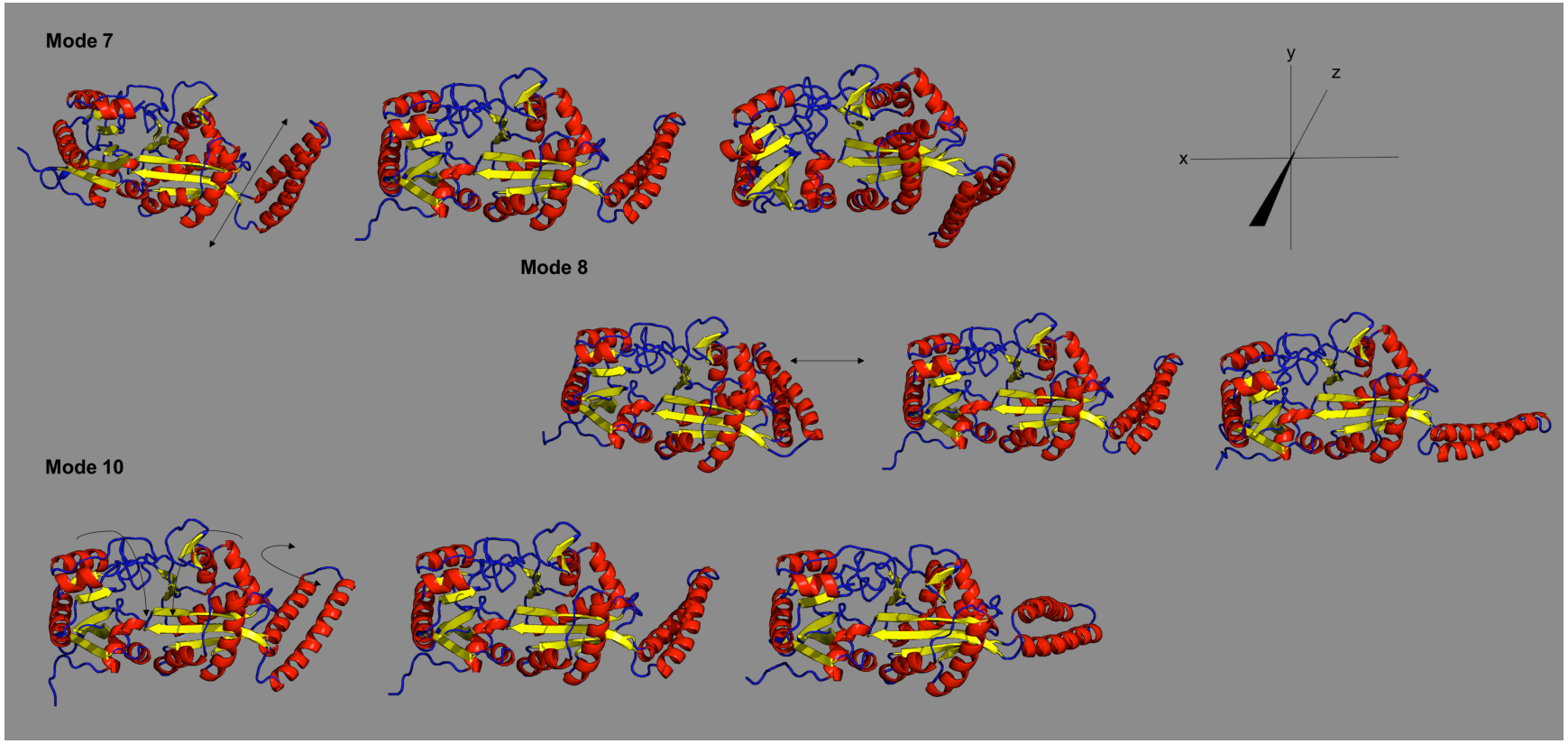



**Figure 5.17: Flexibility Studies of MurM$_{159}$.** Three snapshots of MurM$_{159}$ showing movement in modes 7, 8 and 10 simulated using pdb2movie. Arrows indicate the protein region and direction of movement (Jimenez-Roldan *et al.*, 2012; Römer *et al.*, 2016).

to predict interactions between MurM and possible binding partners.

PG synthesis in streptococci is localised to the septum and at the equatorial rings, with proteins such as PBP1a and PBP2x following FtsZ localization to this region (Scheffers and Pinho, 2005). It is therefore possible that Lipid II(Lys) accumulates in these regions of the bilayer resulting in localisation of MurM. Given that MurM provides the substrate for MurN, GatD/MurT also modify the Lipid II(Lys) and MurJ is required to transport the final lipid through the membrane; it seems unlikely that MurM would act in isolation. MurM may form dimers or multimers with other proteins in this region in order to optimise the final stages of branched PG production. MurM is located in an operon with MurN, which appends the second amino acid to the dipeptide bridge sequentially to MurM. It is possible that the two soluble proteins interact and are transported to the membrane together, or that they dimerise at the membrane in order to channel the Lipid II(Lys) substrate in a similar manner to the GatD/MurT complex which is responsible for amidation of the branched (or unbranched) Lipid II(Lys) substrate (Morlot *et al.*, 2018). Monobranched cross-bridges have been identified in the PG of *S. pneumoniae* MurN deletion mutants (Filipe *et al.*, 2000a), which indicates that the activity of MurM does not require the presence of MurN. However, given that the substrate for MurN is the product of MurM, it stands to reason that MurMN heterodimerisation may assist the activity of MurN by bringing it in close proximity to the membrane and hence its substrate.

Another consideration is that MurM utilises aminoacyl-tRNA as a substrate, MurM may form heterodimers with aaRSs in the cytoplasm, in order to assist the transfer of tRNA to MurM. Furthermore, MurM shares some structural similarity with aaRS (Fiser *et al.*, 2003), some of which have been found to form homodimers (Itoh *et al.*, 2013; Wang *et al.*, 2015), raising the possibility that they may too form heterodimers with proteins such as MurM. In SerRS the coiled-coil helix is required for recognition of the long variable stem of , and tRNA has been shown to span from the coiled-coil arm to the globular domain of this synthetase Biou *et al.* (1994). Whilst MurM

can recognise a tRNA mini-helix as well as full length tRNA (Lloyd *et al.*, 2008), and the *W. viridescens* FemX recognises tRNA in the absence of a coiled-coil region, evidence suggests that the coiled-coil arm of MurM may play an important role in determining amino acid specificity (Filipe *et al.*, 2001b). It is therefore likely that the coiled-coil arm of MurM is involved in the binding of tRNA in a similar manner to SerRS.

Possible sites of MurM interaction for homo- and hetero-dimerisation with MurN and SerRS were investigated using the ZDOCK server (Pierce *et al.*, 2014). In the absence of a *S. pneumoniae* MurN or SerRS, alternative structures or models were used for these predictions. A homology model of MurN was produced using *S. aureus* FemA as a template in I-TASSER (Yang *et al.*, 2015), and the crystal structure of *T. thermophilis* SerRS, which has 37 % sequence identity to *S. pneumoniae* SerRS, was obtained (Zhang, 2008; Roy *et al.*, 2010; Yang *et al.*, 2015; Biou *et al.*, 1994).

All possible binding modes in the translational and rotational space between two proteins is investigated. Any identified pose is then ranked using an energy-based scoring function. The top two predictions for both MurM homodimerisation and heterodimerisation with MurN and *T. thermophilus* SerRS are presented in Figure 5.18.

The two models for MurM homodimerisation both indicate that docking of two MurM proteins would occur at sites, distant from the substrate binding site. This would allow the MurM protein complex to dock to the membrane as shown previously. The first model shows symmetrical docking, such that the sites making contact are the same in both proteins, with regions Ser113-Gln119, Phe379-Pro382 and Ile160-Lys169 being in close proximity. All of these regions are highly conserved among all *S. pneumoniae* strains investigated, with only position 115, 117 and 381 possessing non-conservative mutations. Conversely the second model shows a non-symmetrical binding such that the regions making contact were different for each protein. Whilst this is possible it is more unusual than symmetrical homodimerisation (Goodsell and Olson, 2000).

In the top prediction for heterodimerisation with MurN, the Lipid II(Lys) binding site of MurM is not obstructed by the docking of MurN, thereby allowing membrane interaction and substrate binding. Once MurM has appended the first amino acid, it may undergo a conformational change such as the 'loosening' of the globular domain active site (shown in Mode 10 of Figure 5.17), from around the Lipid II(Lys), allowing dissociation of MurM from both the Lipid II(Lys) and MurN protein. The MurN would then be perfectly poised for binding to the Lipid II(Lys) substrate following addition of L-Ala/L-Ser by MurM. The regions of MurM most likely to form interactions with MurN were identified as Cys111-Gln119, Tyr124-Ile133, Lys140-Ser142 and Met381-Pro385. These positions are mostly highly conserved, however conservative and non-conservative mutations occur between strains at positions 115, 117, 124, 131-133, 140, 381 and 383. In contrast in the second model for MurMN hetero-dimerisation the Lipid II(Lys) binding site is clearly obstructed by the docking of MurN. In this orientation the Lipid II(Lys) would need to fit in the middle of the two proteins in order to gain access to the binding site. Whilst the second model would be feasible, the first model corresponds with the position of MurM at the membrane when Lipid II(Lys) is bound.

Despite being known to form dimers, the structure of *T. thermophilus* SerRS has been solved in the monomeric state and so was used for these heterodimerisation studies. Both models of MurM heterodimerisation with SerRS, show the coiled-coil domains of both proteins coming into close contact with each other, which would permit the successful transfer of aminoacylated-tRNA from the synthetase to the MurM. In the first model, the regions of MurM that are in close proximity to the SerRS are Tyr161-Lys184, Glu378-Met381 and Asn262-Arg266. Whilst Tyr161-Ly184 and Glu378-Met381 are highly conserved (with the exception of positions 175 and 381 which show some variation between strains), Asn252-Arg266 is a highly variable region. In the second model, the regions of MurM in close proximity to the SerRS are Ala158-Lys162, His115-Lys121 and Met381. Ala158-Lys162 are completely conserved and Met381 is highly conserved only showing variation in one strain,

whilst the His115-Lys121 region is highly variable. The presence of highly variable sequences predicted to be present at the protein:protein interface indicates that these interactions may be less likely to occur *in vivo*.

The region Ile160-Lys184 of MurM was identified as a possible site for protein:protein interaction in homodimerisation (first and second models) and heterodimerisation with SerRS (first model). Identification of this region in multiple simulations, in combination with the high sequence conservation across all MurM strains investigated, indicates that this region may be a good candidate for forming important protein:protein interactions. In addition, the region Cys111-Lys121 of MurM was identified as a possible site for protein:protein interaction in heterodimerisation with both SerRS (second model) and MurN (first model), as well as MurM (second model). This region contains a number of highly conserved residues, but also contains a number of residues with high sequence variation. The conserved residues of these regions provide a useful basis for site directed mutagenesis experiments to disturb the protein:protein interactions.

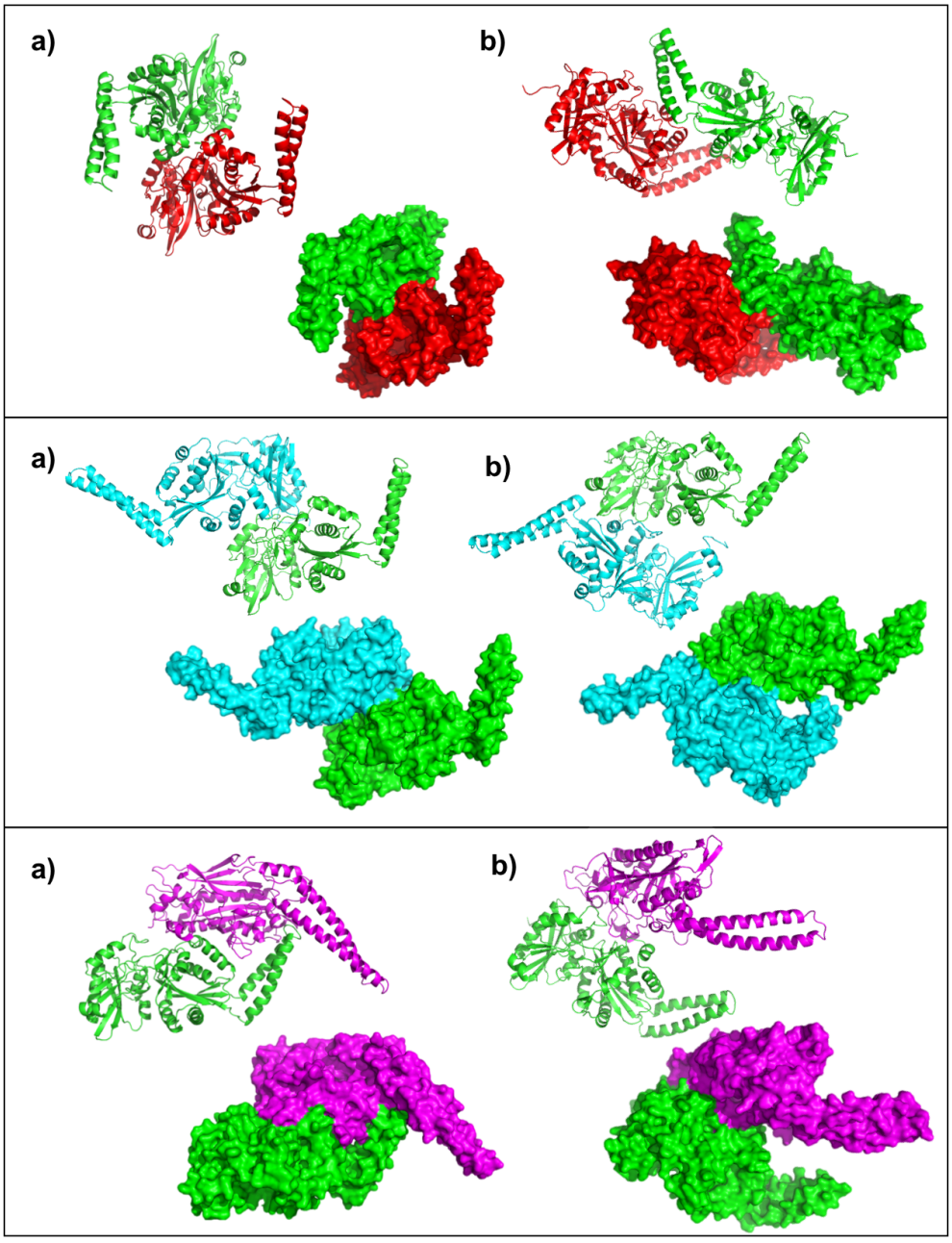



**Figure 5.18: MurM homo- and hetero-dimerisation predictions.** Prediction number 1 (a) and number 2 (b). TOP BOX: homodimerisation of MurM. MIDDLE BOX: heterodimerisation of MurM (green) and MurN (cyan). BOTTOM BOX: heterodimerisation of MurM (green) and SerRS (magenta). Predictions generated using ZDOCK server (Pierce *et al.*, 2014) and visualised in PyMOL (Version 2.2.0).

## 5.8 Discussion

The MurM model presented in this chapter shows a number of improvements when compared to the previous MurM model (Fiser *et al.*, 2003). The new MurM model with an overlaid UDP-MurNAc-5P from *W. viridescens* FemX, resulted in the identification of the Lipid II(Lys) binding site which was not apparent in the previous MurM model (Fiser *et al.*, 2003). This proposed binding site was independently identified in both molecular docking and MD membrane simulations. The orientation of binding of Lipid II(Lys) to the MurM binding site is most likely that identified in the MD membrane simulation, which corresponds well with the orientation of UDP-MurNAc-5P in the *W. viridescens* FemX homologue. The overall orientation of MurM with respect to the membrane when the Lipid II(Lys) is bound, would support the heterodimerisation of MurM with MurN at the membrane. The proximity to the membrane and the orientation of the coiled-coil domain indicate that the flexible nature of this region may be restricted upon MurM binding to the membrane. Furthermore, the binding site is completely occluded by the membrane and so access to the binding site by the tRNA seems ever more unlikely. In addition, the composition of phospholipids in the membrane may modulate both association of MurM to the membrane and its catalytic activity. Taken together these simulations support the hypothesis that MurM may use a ping-pong mechanism of action. In the cytoplasm the amino acid would be transferred from the tRNA to the MurM to create an aminoacylated-MurM. MurM then travels to the membrane, binds Lipid II(Lys) and appends the amino acid to the Lipid II(Lys) head-group. MurM may also form protein:protein interactions with a range of cytoplasmic and membrane proteins in order to channel substrates effectively.

Atomistic MD simulations of MurM with the membrane have provided novel insight into how and where Lipid II(Lys) interaction with MurM occurs. In addition these simulations have shown how phospholipids and MurM form interactions which may modulate membrane association or activity of MurM. In the simulation

containing no CL, MurM successfully associated with the membrane indicating that CL is not required for this interaction. However, in simulations containing CL, the CL was highly enriched in the area of membrane making contact with MurM, indicating that interactions with MurM are important. It remains unknown as to whether *in vitro* MurM recruits CL or whether the CL localises in the membrane forming a cluster which recruits MurM. It is important to consider that the membrane of bacteria *in vivo* is vast and highly dynamic and that these simulations use a simplistic membrane which does not accurately represent the membrane composition of *S. pneumoniae*.

Flexibility modelling indicates that the coiled-coil arm is highly flexible, which is likely a contributory factor to the previous lack of success when attempting to crystallise this protein. Stabilising this domain by addition of tRNA may eliminate these large scale motions and assist the formation of crystals. Smaller scale motions observed within the globular domain may be indicative of a conformational change that occurs upon Lipid II(Lys) binding. Whilst these simulations are useful in identifying regions of space which are permissible to the molecules, it is important to note that this algorithm provides information of the physical flexibility of this protein in the absence of water or any other forces acting on it. These predicted movements therefore represent regions of MurM which show flexibility but may not necessarily be of biological significance or occur *in vivo*. Whilst this method is not as accurate as atomistic modelling such as MD it remains a useful tool due to the short run time and the good predictive quality. The information garnered from flexibility studies of MurM can be used to identify distant regions on the protein that may make contact upon conformational change.

ZDOCK predictions of MurM homo- and hetero-dimerisation identified potential sites of protein:protein interactions indicating that MurM may not only form interactions with the membrane but also with other partner proteins. Further work to identify binding partners of MurM, and better characterise the protein:protein interface are required.

In conclusion, Box and Draper (1987) once said "all models are wrong, but some are useful"; in this chapter, modelling and simulation has proved to be highly useful. In the absence of a MurM crystal structure, modelling and simulation has provided insight into not only the structure of the MurM protein itself, but also the possible movements of this protein and potential interactions it may make with the membrane and other proteins. When taken together this information provides additional insight into MurM's mechanism of action which will help direct and focus future experimental research.

## 5.9 Future work

Docking experiments in the absence of the hydrocarbon lipid tail and diphosphates, should be conducted in order to identify the substrate orientation and important residues involved in binding of Lipid II(Lys). Subsequently, MD using the Nudged Elastic Band (NEB) method may be used in order to identify the low energy transition pathway of substrate binding to MurM.

Residues of importance could be confirmed experimentally by single and combinatorial mutations of residues in the MurM binding site, which were proposed to be important for Lipid II(Lys) binding. In order to establish residues important for binding, the dissociation constant ($k_d$) of Lipid II(Lys) and MurM could be compared for MurM mutants using microscale thermophoresis (MST). In addition, the effects of different mutations of the catalytic activity of MurM could be investigated using a MurM assay. Since the MurM residue at position 37 is a conserved Asn, in all penicillin sensitive strains evaluated, and a conserved D, in all penicillin resistant strains evaluated, reciprocal mutation of this residue may be informative for both effect on the catalytic efficiency of MurM *in vitro* and any effect on penicillin resistance, or alterations in the PG composition *in vivo*.

Further simulation and analysis of MurM at the membrane may identify additional residues directly involved in MurM:cardiolipin interactions. The effect of mutating

these residues may be investigated experimentally using the MurM assay in the presence of cardiolipin.

Possible MurM interactions with other proteins present at the membrane, such as MurJ or MurT/GatD should also be predicted using ZDOCK. Experimental investigation of MurM/MurN heterodimerisation has previously been hindered by difficulty in expressing and purifying MurN. Previous work suggests that MurN co-purifies with the Maltose Binding Protein (MBP) affinity tag, and cannot be separated, supporting the idea that MurN may form protein:protein interactions. The presence of this MBP-tag may however interfere with the *in vitro* experimental techniques proposed, and so, since MurM and MurN exist in an operon, it may be beneficial to overexpress them together in pET28a rather than individually. Native-PAGE, pull-down assays, analytical ultracentrifugation (AUC) or *in vivo* or *in vitro* cross-linking experiments may confirm dimerisation between the suggested proteins. Carbene labelling is a technique that may be exploited to experimentally confirm the Lipid II(Lys) binding site (but not substrate orientation) and confirm other possible protein:protein interactions. Subsequent biomolecular interaction techniques such as isothermal calorimetry (ITC), surface plasmon resonance (SPR) or MST could be used to establish binding kinetics. The radiolabelled MurM assay may be utilised to explore binding kinetics of MurM with the Lipid II(Lys) substrate. If MurM has a ping-pong mechanism of action then Lineweaver-Burk plots at varying Lipid II(Lys) concentrations would occur as parallel lines.

Following experimental identification of MurM homo- or hetero-dimerisation, site-directed mutagenesis, guided by these models, may be conducted to disturb the protein:protein interactions and identify specific residues that are important for protein:protein interactions.

# Chapter 6

# Final discussion and future perspectives

The main aim of this work was to characterise the effects of $H_2O_2$ on the pre- and post-transfer editing activity of AlaRS and ThrRS, and to identify specific modifications arising due to $H_2O_2$. Secondary aims were to use a combination of experimental and computational approaches to gain a better understanding of MurM, its substrate specificty and its mechanism of action.

## 6.1 Characterising the effect of $H_2O_2$ on aminoacyl-tRNA synthetases

Building on previous observations by Ling and Söll (2010), that oxidation by $H_2O_2$ resulted in an increase in production of seryl-tRNA$^{Thr}$ by *E. coli* ThrRS, the effects of $H_2O_2$ on *S. pneumoniae* AlaRS and ThrRS were characterised. This work utilised a novel spectrophotometric editing assay which enabled the real-time monitoring of pre- and post-transfer editing *in vitro*. This work supports previous findings, and demonstrates that $H_2O_2$ specifically affects post-transfer editing, but not pre-transfer editing of both AlaRS and ThrRS. It was also shown that oxidation of AlaRS and

ThrRS is dependent upon the relative proportions of $H_2O_2$ and aaRS, as well as time. Initial concentrations of $H_2O_2$ were varied from 0 mM to 75 mM demonstrating a decrease in post-transfer editing activity. Incubation of a higher concentration of enzyme, with a concentration of $H_2O_2$ previously identified to cause a reduction in post-transfer editing activity, did not generate any loss in activity of AlaRS or ThrRS. This result indicated that, as expected oxidation by $H_2O_2$ is a bi-molecular reaction and so it was not the concentration of $H_2O_2$ per se which was important, but rather the molar ratio of enzyme to $H_2O_2$. As such incubation experiments were conducted with $H_2O_2$ at 10,000-fold and 200-fold $H_2O_2$ to AlaRS or ThrRS respectively. Whilst from these ratios it is clear that ThrRS is much more sensitive to $H_2O_2$ than AlaRS, these ratios seem reasonable given that previous experiments by Ling and Söll (2010) on ThrRS were conducted at a 2666-fold excess of $H_2O_2$ (corresponding to 4 mM $H_2O_2$ in the assay with 1.5 $\mu$M ThrRS). Due to the high concentrations of tRNA required in this novel editing assay it was not possible to characterise differences in editing between individual tRNA isoacceptors. However, individual tRNA isoacceptors are suitable for use in radiolabelled MurM assays and can be used to investigate any difference in use as MurM substrates.

Demonstrating that exposure to $H_2O_2$ results in a loss of post-transfer editing in AlaRS and ThrRS, prompted experiments to identify specific modifications which corresponded to this loss in activity. Cys182 of *E. coli* ThrRS was identified as the residue targeted by oxidation with $H_2O_2$ by experiments using Ellman's reagent (also known as 5,5'-dithiobis-(2-nitrobenzoic acid); DTNB) and subsequently DAz-2 (Wu *et al.*, 2014). Whilst this does indicate that Cys182 is oxidised, it does not confirm if any other (non-cysteine) residues such as methionine or tryptophan are oxidised. Furthermore DAz-2 only identifies sulphenic acid and so if cysteine is oxidised rapidly to sulphonic acid, the experiment would falsely indicate that there was no oxidation occurring.

In order to keep these experiments consistent with the loss of post-transfer editing function, the results from previous experiments were used to inform the incubation

time and relative ratio of $H_2O_2$ required for aaRSs treatment. Intact protein mass spectrometry suggested that there was a shift in mass between untreated and treated samples, however the poor quality data, which was due to experimental limitations, meant that further investigation was required. In an attempt to identify individual oxidised residues and the oxidation state, peptide-based mass spectrometry was used.

Initial proteomics data suggested that this protein was highly oxidised, AlaRS and ThrRS showed oxidation at a total of 19 and 26 residues respectively. The single Cys669 of AlaRS was oxidised to sulphonic acid in both treated and the untreated control. This indicates that either the experimental procedure introduced oxidation to the control sample, or that the modifications which result in the loss of post-transfer editing activity are more complex than presence or absence of oxidation and may instead pertain to the proportion of sample with oxidised residues. The mass spectrometry techniques used here were unable to provide any quantitative data and so it is possible that only a small proportion of control sample was oxidised, compared to a larger proportion in the treated sample. Therefore, improved mass spectrometry experiments, which minimise additional preparative oxidation and allow quantitative analysis of oxidation must be used to determine residues which, when oxidised by $H_2O_2$, result in a loss of post-transfer editing of both AlaRS and ThrRS. Since cysteine residues have previously been demonstrated to be important for post-transfer editing, a DTNB assay could be used to monitor oxidation of these residues over time.

*E. coli* ProRS contains a Cys443; whilst a Gly443 mutant can still conduct post-transfer editing, it is severely defective suggesting that Cys443 has an important role in the editing mechanism (Beuning and Musier-Forsyth, 2000). The homologous residue in *S. pneumoniae* ProRS is a Cys477 (the only cysteine in the open reading frame), therefore it would be interesting to repeat these experiments to investigate whether $H_2O_2$ effects the editing activity of ProRS in a similar way to that of AlaRS and ThrRS. Interestingly, ProRS has a *trans*-editing factor YbaK, which does not

contain Cys477. This suggests that YbaK has an additional mechanism for editing which does not require a single cysteine. In *S. pneumoniae*, if oxidation of Cys477 by $H_2O_2$ resulted in a reduction of post-transfer editing activity, the YbaK homologue would likely provide protection by editing mis-charged tRNA using this alternative mechanism, hence maintaining the fidelity of protein synthesis.

Now that the effects of $H_2O_2$ on aaRSs have been characterised *in vitro*, it would be prudent to investigate the biological relevance *in vivo*. *S. pneumoniae (159)* has been shown to preferentially incorporate alanine into the PG when grown in the presence of catalase (mimicking anaerobic conditions), and $MurM_{159}$ demonstrated a preference towards alanyl-$tRNA^{Ala}$ when compared to seryl-$tRNA^{Ser}$ *in vitro* (Lloyd *et al.*, 2008). However, Shepherd (2011) subsequently demonstrated that *in vitro* $MurM_{159}$ utilised mis-charged seryl-$tRNA^{Ala}$ preferentially. Therefore, if MurM directs mis-charged tRNA into the PG under aerobic growth conditions, in order to maintain the fidelity of protein synthesis, it would be expected that the composition of the PG would vary accordingly. *S. pneumoniae (159)* would therefore be a good strain to use for *in vivo* experiments. Comparison of the PG from wild-type *S. pneumoniae (159)* and *S. pneumoniae (159)*ΔSpxB, grown in aerobic conditions, would identify any changes in PG composition which can be attributed to the presence or absence of $H_2O_2$. Since aerobic growth of *S. pneumoniae* is challenging, the same result may be demonstrated by growth of *S. pneumoniae* in anaerobic conditions with the addition of exogenous $H_2O_2$ into the growth media. Additionally, whilst creating Cys669 mutants for AlaRS or Cys181 and Cys336 mutants for ThrRS, would most likely be lethal, transformation of *S. pneumoniae*, grown anaerobically, with donor DNA containing mutant aaRS may result in an increase in mis-charging (mimicking oxidative stress) but maintain mis-charging below the lethal threshold. The increase in mis-charged tRNA would be expected to alter the composition of the PG accordingly. The composition of PG can be analysed by mass spectrometry, and so any increase in serine incorporation due to aerobic growth, the presence of exogenous $H_2O_2$ or the introduction of mutant AlaRS or ThrRS can be monitored

(Severin and Tomasz, 1996; Garcia-Bustos and Tomasz, 1990; Lloyd *et al.*, 2008).

## 6.2 Characterising the substrate specificity of MurM

The sequence of $MurM_{159}$ and $MurM_{Pn16}$ was clarified, and due to their utility for comparison between sensitive and highly penicillin resistant strains, full genome sequencing of each was conducted. This provides the basis for other proteins of interest, such as PBPs, MurN and GatD/MurT from these strains to be investigated further.

The presence of iGln Lipid II(Lys) in the PG of *S. pneumoniae* and the recent identification of the MurT/GatD complex responsible for amidation, led to the question of the order that these modifications occur to the Lipid II(Lys) substrate. It is possible that there is a strict order to amidation and branching of Lipid II(Lys), or that there is no order, such that MurT/GatD can amidate both branched or unbranched Lipid II(Lys) and MurM can branch both amidated or unamidated Lipid II(Lys). Work by Morlot *et al.* (2018) investigating depletion of MurT/GatD showed that, upon depletion of this complex, as expected, the proportion of amidated PG, and cross-linked PG decreased. However, unexpectedly it was observed that there was a concomitant increase in branched PG, and so was suggested that Lipid II(Lys) may be a better substrate for MurM (Morlot *et al.*, 2018). All previous work with MurM has been conducted with Lipid II(Lys). This is because synthesis of iGln Lipid II(Lys) results in very poor yields and this has been the limiting factor to investigating iGln Lipid II(Lys) as a MurM substrate. In this work we improved the yield of iGln Lipid II(Lys) and subsequently showed that both Lipid II(Lys) and iGln Lipid II(Lys) are substrates for $MurM_{159}$ *in vitro*. Future kinetic analysis of Lipid II(Lys) and iGln Lipid II(Lys) would demonstrate whether MurM utilises these substrates equally and therefore shed light onto any order to Lipid II(Lys) modifications *in vivo*.

The absence of a G at position 51 or 63 of tRNA has been identified as a negative

identity element for EF-Tu. It has been reported that in *T. thermophilus*, *E. coli* and *S. aureus* tRNA without a G at this position bind EF-Tu weakly. It has been proposed that tRNAs containing this negative identity element are retained for non-proteogenic functions such as PG biosynthesis. As such it is important to characterise the ability of MurM to utilise different tRNA isoacceptors as substrate. The work in this thesis paves the way to conduct radioactive MurM experiments for characterisation of different tRNA isoacceptors charged with cognate and non-cognate amino acids. Crude *S. pneumoniae (159)* tRNA was prepared, cognate charging by AlaRS, SerRS and ThrRS was demonstrated in a radiolabelled assay and a good yield of charged tRNA was purified during a preparative charging assay. In addition, $tRNA^{Ala}$, $tRNA^{Ser1}$, $tRNA^{Ser2}$, $tRNA^{Ser3}$, $tRNA^{Thr1}$ and $tRNA^{Thr2}$ were synthesised by *in vitro* transcription. All $tRNA^{Ala}$ and $tRNA^{Ser}$ isoacceptors demonstrated cognate charging by AlaRS and SerRS (3 $\mu$M) in a radiolabelled assay. No charging of $tRNA^{Thr}$ isoacceptors was observed by ThrRS at the same concentration; since ThrRS was found to be significantly less active than AlaRS in both amino acid activation and editing assays, the low activity of the enzyme may explain the lack of charging of the $tRNA^{Thr}$ isoacceptors, and so this may be overcome with a higher concentration of ThrRS in charging assays.

Some preliminary experiments generating radiolabelled mis-charged tRNA, indicated that addition of high concentrations of $H_2O_2$ (aimed at abolishing post-transfer editing) completely eliminated tRNA charging. This suggests that at such high concentrations of $H_2O_2$ other components of the reaction such as the tRNA is adversely effected. Therefore, the proposed best method to obtaining mis-charged tRNA would be to treat the aaRS with $H_2O_2$, dialyse the $H_2O_2$ and subsequently use the aaRS in the radiolabelled assay. Alternatively, for AlaRS and ThrRS, editing site mutants can be used to generate large quantities of mis-charged tRNA for radiolabelled experiments.

Once tRNAs charged with cognate and non-cognate amino acids have been generated, any preference of MurM to use different $tRNA^{Ser}$ and $tRNA^{Thr}$

isoacceptors can be investigated using radioactive MurM activity assays.

*S. aureus* contains several $tRNA^{Gly}$ isoacceptors which do not posses the G at this position and so are thought to be reserved for PG synthesis. Similarly, $tRNA^{Ser1}$, $tRNA^{Ser2}$ and $tRNA^{Thr2}$ do not posses G at this position and are likely to be EF-Tu unstable. Therefore it would be interesting to compare the use of $tRNA^{Ser1}$, $tRNA^{Ser2}$ and $tRNA^{Ser3}$, and $tRNA^{Thr1}$ and $tRNA^{Thr2}$ as MurM substrates. $tRNA^{Ser2}$ is a known MurM substate since all previous published experiments were conducted with this isoacceptor. In addition the 51:63 bp is located in the T$\psi$C loop, which is required in the $tRNA^{Ala}$ minihelix which is also a substrate for MurM. It is therefore possible that the absence of a G at this position may also double as a positive identity element for MurM recognition, and so introduction of a G at this position in a non-proteogenic isoacceptor may enhance substrate recognition by MurM.

## 6.3 An improved understanding of the structure of MurM

A variety of computational techniques were used to model the structure of MurM and predict its interactions with the Lipid II(Lys) substrate, the cytoplasmic face of the membrane and other proteins. Despite using the same modelling software, some significant differences were observed between the previous model of MurM (Fiser *et al.*, 2003) and the model generated in this work. The most likely explanation for these differences is the homologue chosen as a template for the model; the previous MurM model was generated using the crystal structure of *S. aureus* FemA, whilst the new model detailed in this thesis used an unpublished structure of *S. aureus* FemX (unpublished work from Roper lab). FemX, similarly to MurM, is responsible for the addition of the first amino acid in the peptide bridge and so is a functional homologue. In addition, the sequence identity between MurM and FemX is higher than with FemA.

The new structure of MurM allowed identification of a MurM binding site which contradicts the site previously proposed. In addition, the groove which was

previously proposed to be the Lipid II(Lys) binding site was not observed in the new model, due to differences between the tertiary structure of the two models. The prediction of the binding site in the previous MurM model was based on the structure-function analogy between MurM and NMT proteins (Fiser *et al.*, 2003). However, whilst NMT binds lipid, MurM only binds the pentapeptide headgroup of Lipid II, the prenyl tail of which is embedded in the membrane. The MurM Lipid II(Lys) is more similar to the substrate of *W. viridescens* FemX, despite being a soluble (and not Lipid-linked) PG precursor, than that of an NMT. When MurM was compared with *W. viridescens* FemX, the binding sites overlaid well, moreover the substrate of *W. viridescens* FemX appeared to fit into the proposed MurM binding site. Since the MurM substrate is membrane embedded Lipid II(Lys), molecular docking experiments were conducted with a truncated version which did not possess the lipid tail. The same binding site was independently identified by molecular docking studies, supporting this as the true MurM binding site, however the orientation of the truncated Lipid II(Lys) in the binding site made little biological sense. MD simulations were used to investigate the interactions of MurM with the membrane. Unexpectedly perhaps, these simulations also identified an interaction between Lipid II(Lys) and the newly proposed binding site of MurM, further supporting this as the Lipid II(Lys) binding site of MurM. From the interaction of MurM with Lipid II(Lys) embedded in the membrane, a number of conserved residues involved in hydrogen bonding were identified. These studies were able to demonstrate association of MurM to the membrane, and also the effect of MurM:membrane interactions between phospholipids such as PhG and CL. Interplay between MurM and the membrane phospholipids supports previous experimental evidence which suggests that phospholipids in the membrane regulate MurM's activity.

MurM was predicted to form homo- and hetero- dimers with a number of proteins including MurN. This supports the possibility that MurM and MurN are targeted to the membrane together, in order to streamline the sequential addition of amino acids to form the dipeptide bridge. Furthermore, the orientation of MurM bound to

Lipid II(Lys) at the membrane, exposes the surface of MurM proposed to interact with MurN and so would permit heterodimerisation with MurN whilst bound to Lipid II(Lys).

The orientation of MurM interacting with Lipid II(Lys) at the membrane shows that the flexible coiled-coil domain is in close proximity to the membrane. Therefore, given that this domain is proposed to be important for tRNA binding, it is unlikely that MurM with tRNA bound could adopt this position at the membrane. It is therefore proposed that the most likely mechanism of action of MurM is the ping-pong mechanism, whereby MurM undergoes an acylation reaction in the cytoplasm, and subsequently travels to the membrane with the amino acid for transfer to Lipid II(Lys).

With the exception of residues important for the amino acid specificity (within the coiled-coil domain) and the conserved Asp107 proposed to be important for catalysis, these studies identify, for the first time, key residues in the MurM binding site. Experiments conducted with both individual and combined site-directed mutations at these residues, would likely reveal the importance of these residues on binding and catalysis. In addition, further analysis of MurM:membrane interactions may identify additional residues important for phospholipid interactions. Finally predictions of MurM homo- and hetero-dimerisation indicates surfaces of MurM which may be involved in protein:protein interactions and so could also be investigated by *in vivo* mutagenesis experiments.

In conclusion, this work has generated data important in demonstrating that $H_2O_2$ causes a loss of post-transfer editing in multiple *S. pneumoniae* aaRSs *in vitro*. Furthermore, significant progress has been made in preparation for investigating the role of MurM in maintaining the fidelity of protein synthesis. Future experiments proposed will provide *in vivo* evidence demonstrating a link between protein synthesis and PG synthesis. Finally, computational studies have enhanced our understanding of the structure and interactions of MurM, which together shed light onto this protein's possible mechanism of action. These studies will help guide *in*

*vitro* and *in vivo* experiments which can further elucidate binding partners, substrate specificity and mechanism of action for MurM. Overall, having an enhanced structural and functional understanding of MurM, makes this protein a good prospective target for the development of novel antimicrobial resistance inhibitors which, by inhibition of MurM, could restore penicillin sensitivity in highly resistant strains of *S. pneumoniae*.

# Chapter 7

# Appendix

## 7.1 Gene fragments and primers for cloning

| **Construct** | **Target** | **Primer Sequence 5′-3′** |
|---|---|---|
| pET28::ThrRS | Plasmid | GGGAAAGTAATGTTAATCATGCTGT GATGATGATGATGATGGC |
| | | AATCACGCGTTGAGAAATAATGGGT CGCGGATCCGAATTC |
| | Insert | ATGATTAACATTACTTTCCCAGATG |
| | | TTATTTCTCAACGCGTGATTTG |
| pET22b::gMurM$_{159}$ | Plasmid | ATGTATATCTCCTCCTTAAAGTTAA ACAAAATTATTTCTAGAGGGGAATT GTTATCCGC |
| | | TGAGATCCGGCTGCTAAC |
| | Insert | TTGTTAGCAGCCGGATCTCATCAGT GGTGGTGGTGGTG |
| | | TTTAAGAAGGAGATATACATATGTA CAGATACCAGCTTGGAATC |

| | | |
|---|---|---|
| pET22b::gMurM$_{Pn16}$ | Plasmid | ATGTATATCTCCTTCTTAAAGTTAA ACAAAATTATTTCTAGAGGGGAATT GTTATCCGC |
| | | TGAGATCCGGCTGCTAAC |
| | Insert | TTGTTAGCAGCCGGATCTCATCAGT GGTGGTGGTGGTG |
| | | TTTAAGAAGGAGATATACATATGTA TCGCTATCAAATCGGTATCC |

**Table 7.2: Primers used for cloning using the Gibson assembly® method.** The final construct name and forward and reverse primers, for both plasmid and insert, are shown.

## 7.2 Construct details

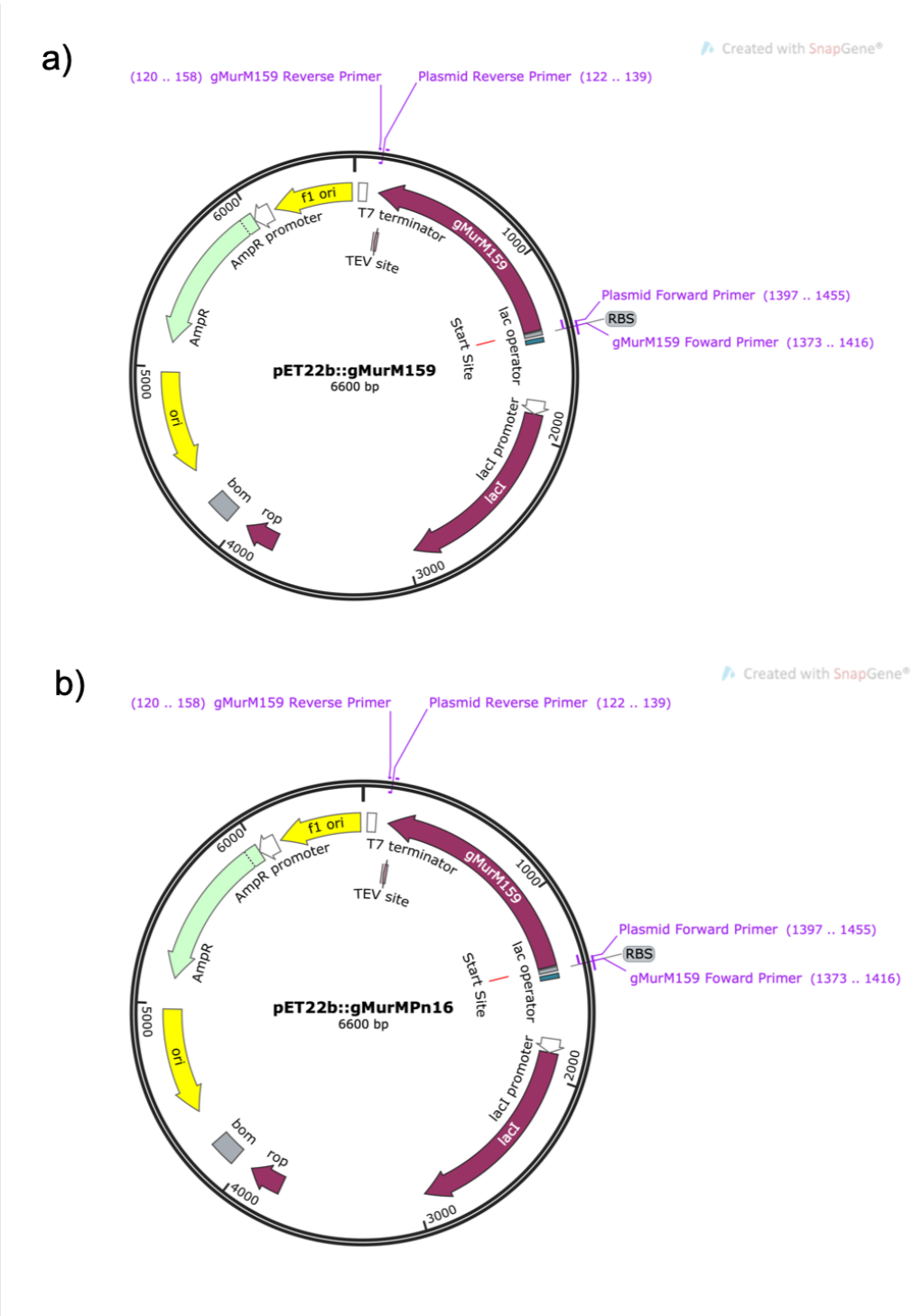



**Figure 7.1: Construct maps for MurM.** a) $gMurM_{159}$ in pET22b(+) cloned using the Gibson assembly® method. The binding sites for plasmid and insert primers are indicated. b) $gMurM_{Pn16}$ in pET22b(+) cloned using the Gibson assembly® method. The binding sites for plasmid and insert amplification primers are indicated. Figure made using GeneSnap®.

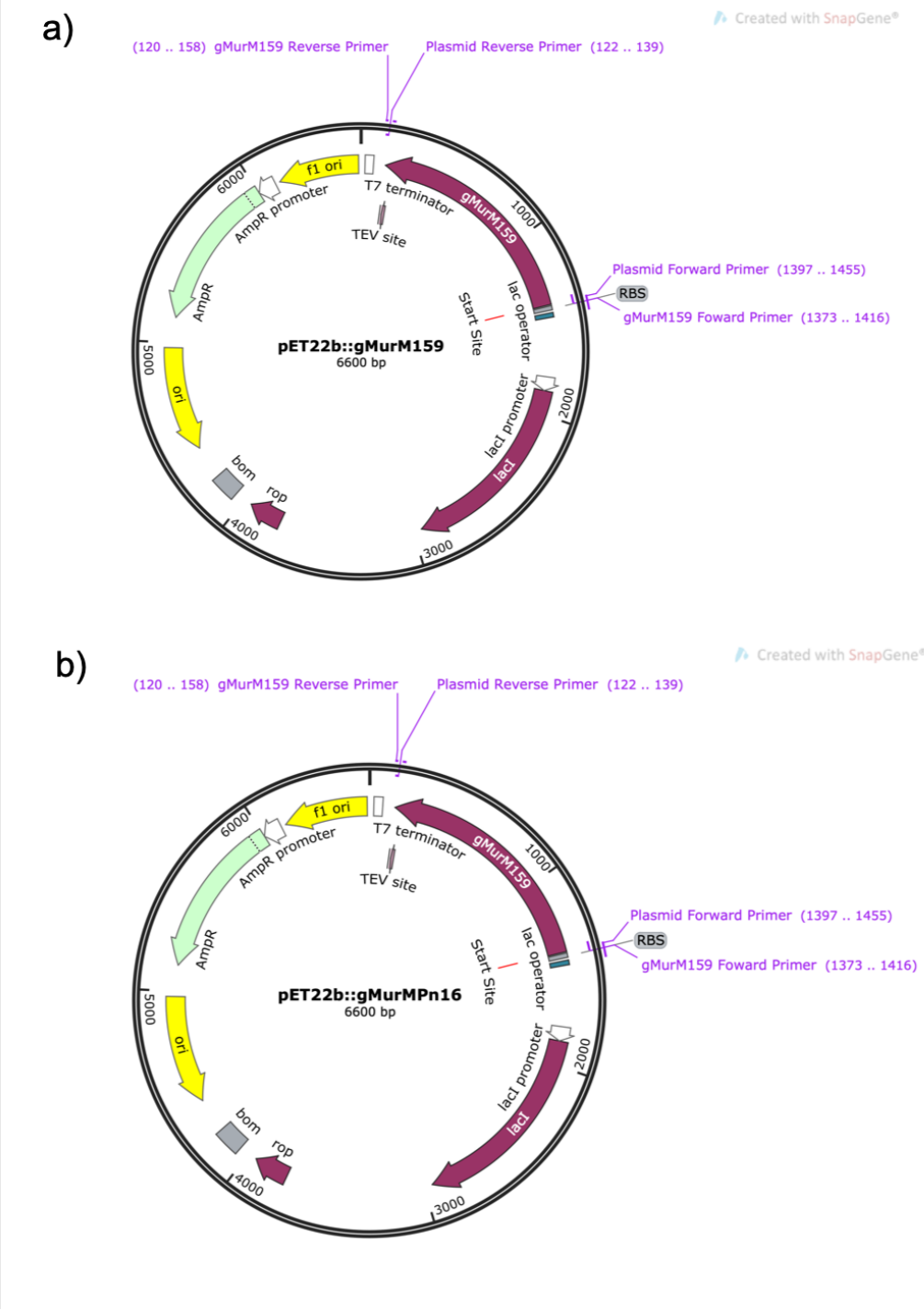



**Figure 7.2: Construct map for ThrRS.** ThrRS(N-term His-tag) in pET28a(+) cloned using the Gibson assembly® method. The binding sites for plasmid and insert amplification primers are indicated. Figure made using GeneSnap®.

## 7.3 Protein sequences

a)

```
        10         20         30         40         50         60
MYRYQLGIPL SEYDGFVKEH PMVNLLQSSA WEKVKSDWNH ERLGVYEGEN LLAVASILIK

        70         80         90        100        110        120
SLPLGYKMFY IPRGPILDYR DTELLKFVLQ SIKSYARSKR AVFVTFDPSI CLSQHLVNQD

       130        140        150        160        170        180
KREYPENLAI VEILGQLGVK WSGRTIEMDD TIQPRIQAKI YKENFEEDKL SKSTRQAIRT

       190        200        210        220        230        240
ARNKGLEIQY GGLELLDSFS ELMKKTEKRK EIHLRNEAYY KKLLDNFKED SYITLTSLDV

       250        260        270        280        290        300
SKRLRELEEQ LEKNRVVAEK FNDATRSSKV QENIKEKERL KEEIDFLQGY MNMGKSNIPL

       310        320        330        340        350        360
AATLSLEFGN TSVNLYAGMD DDFKRYNAPI LTWYETARYA FERGMVWQNL GGVENSLNGG

       370        380        390        400
LYQFKEKFNP TIEEYLGEFT MPTHPLYPLL RLALDFRKTL RKKHRK
```

Theoretical pI/Mw: 8.81 / 47516.49

b)

```
        10         20         30         40         50         60
MYRYQIGIPT LEYDQFVKEH ELANVLQSSA WEEVKSNWQH EKFGVYREEK LLATASILIR

        70         80         90        100        110        120
TLPLGYKMFY IPRGPILDYG DKELLNFAIQ SIKSYARSKR AVFVTFDPSI CLSQSLINQE

       130        140        150        160        170        180
KTEFPENLAI IDSLQQMGVR WSGKTEEMGD TIQPRIQAKI YKENFEEDKL SKSTKQAIRT

       190        200        210        220        230        240
ARNKGLEIQY GGLELLDSFS ELMKKTEKRK EIHLRNEAYY KKLLDNFKDK AYITLATLDV

       250        260        270        280        290        300
SKRSQELEEQ LAKNRALEET FTESTRTSKV EAQKKEKERL LEELTFLQEY IDVGQARVPL

       310        320        330        340        350        360
AATLSLEFGT TSVNIYAGMD DDFKRYNAPI LTWYETARYA FERGMVWQNL GGVENSLNGG

       370        380        390        400
LYHFKEKFNP TIEEYLGEFT MPTHPLYPLL RLALDFRKTL RKKHRK
```

Theoretical pI/Mw: 8.22 / 47413.20

**Figure 7.3: MurM protein sequences.** a) $MurM_{159}$ and b) $MurM_{Pn16}$.

```
        10          20          30          40          50          60
MKQLSSAQVR QMWLDFWATK GHSVEPSVSL VPVNDPTLLW INSGVATLKK YFDGTIIPEN

        70          80          90         100         110         120
PRITNAQKAI RTNDIENVGK TARHHTMFEM LGNFSIGDYF RDEAITWAYE LLTSPEWFDF

       130         140         150         160         170         180
PAEKLYMTYY PDDKDSYNRW IEVGVDPSHL IPIEDNFWEI GAGPSGPDTE IFFDRGEAFD

       190         200         210         220         230         240
PENIGLRLLA EDIENDRYIE IWNIVLSQFN ADPAVPRSEY KELPHKNIDT GAGLERLVAV

       250         260         270         280         290         300
IQGAKTNFET DLFMPIIREV EKLSGKVYDQ DGDNMSFKVI ADHIRSLSFA IGDGALPGNE

       310         320         330         340         350         360
GRGYVLRRLL RRASMHGQKL GINEPFLYKL VPTVGKIMES YYPEVLEKRD FIEKIVKSEE

       370         380         390         400         410         420
ESFARTLHSG QHFAQGIVAD LKEKGQSVIA GQDVFKLYDT YGFPVELTEE IAEEAGMTVD

       430         440         450         460         470         480
REGFEAAMKE QQERARASAV KGGSMGMQNE TLQNITVESV FNHNASQLSS KLVAIVADNA

       490         500         510         520         530         540
EVEAVSEGTT SLIFAETPFY AEMGGQVADH GQILDESGKV VATVTNVQKA PNGQALHTVE

       550         560         570         580         590         600
VLAPLALNQE YTLAIDSNRR HRVMKNHTAT HLLHAALHNI LGNHATQAGS LNEVEFLRFD

       610         620         630         640         650         660
FTHFQAVTAE ELRAIEQQVN EKIWEALEVK TVETDIDTAK EMGAMALFGE KYGKEVRVVT

       670         680         690         700         710         720
IGDYSIELCG GTHVGNTSEI GLFKIVKEEG IGSGTRRILA VTGKEAFEAY REQEDALKAV

       730         740         750         760         770         780
AATLKAPQVK EVPHKVEGLQ EQLRQLQKEN AELKEKAAAA AAGDIFKDVK EVNGHRYIAS

       790         800         810         820         830         840
QVSVSDAGAL RTFADNWKQK DYSDLLVLVA AIGDKVNVLV ASKTKDLHAG NLVKELAPII

       850         860         870         880
DGRGGGKPDM AMAGGSNQPK IQELLDAVAG KLLEHHHHHH
```

Theoretical pI/Mw: 5.19 / 97534.11

**Figure 7.4: AlaRS protein sequence.** Theoretical pI and molecular weight (MW) were calculated using ExPASy (Artimo *et al.*, 2012)

```
        10         20         30         40         50         60
MINITFPDGA VREFESGVTT FEIAQSISNS LAKKALAGKF NGKLIDTTRA ITEDGSIEIV

        70         80         90        100        110        120
TPDHEDALPI LRHSATHLFA QAARRLFPDI HLGVGPAIED GFYYDTDHTA GQISNEDLPR

       130        140        150        160        170        180
IEEEMQKIVK ENFPSIREEV TKDEAREIFK NDPYKLELIE EHSEDEGGLT IYRQGEYVDL

       190        200        210        220        230        240
CRGPHVPSTG RIQIFHLLHV AGAYWRGNSD NAMMQRIYGT AWFDKKDLKN YLQMREEAKE

       250        260        270        280        290        300
RDHRKLGKEL DLFMISQEVG QGLPFWLPNG ATIRRELERY IVNKELVSGY QHVYTPPLAS

       310        320        330        340        350        360
VELYKTSGHW DHYQEDMFPT MDMGDGEEFV LRPMNCPHHI QVFKHHVHSY RELPIRIAEI

       370        380        390        400        410        420
GMMHRYEKSG ALTGLQRVRE MSLNDGHLFV TPEQIQEEFQ RALQLIIDVY EDFNLTDYRF

       430        440        450        460        470        480
RLSLRDPQDT HKYFDNDEMW ENAQTMLRAA LDEMGVDYFE AEGEAAFYGP KLDIQIKTAL

       490        500        510        520        530        540
GKEETLSTIQ LDFLLPERFD LKYIGADGED HRPVMIHRGV ISTMERFTAI LIENYKGAFP

       550        560        570        580        590        600
TWLAPHQVTL IPVSNEKHVD YAWEVAKKLR DRGVRADVDE RNEKMQFKIR ASQTSKIPYQ

       610        620        630        640
LIVGDKEMED ETVNVRRYGQ KETQTVSVDN FVQAILADIA NKSRVEK

Theoretical pI/Mw: 5.26 / 74764.46
```

**Figure 7.5: ThrRS protein sequence.** Theoretical pI and molecular weight (MW) were calculated using ExPASy (Artimo *et al.*, 2012)

## 7.4 Tandem mass spectra

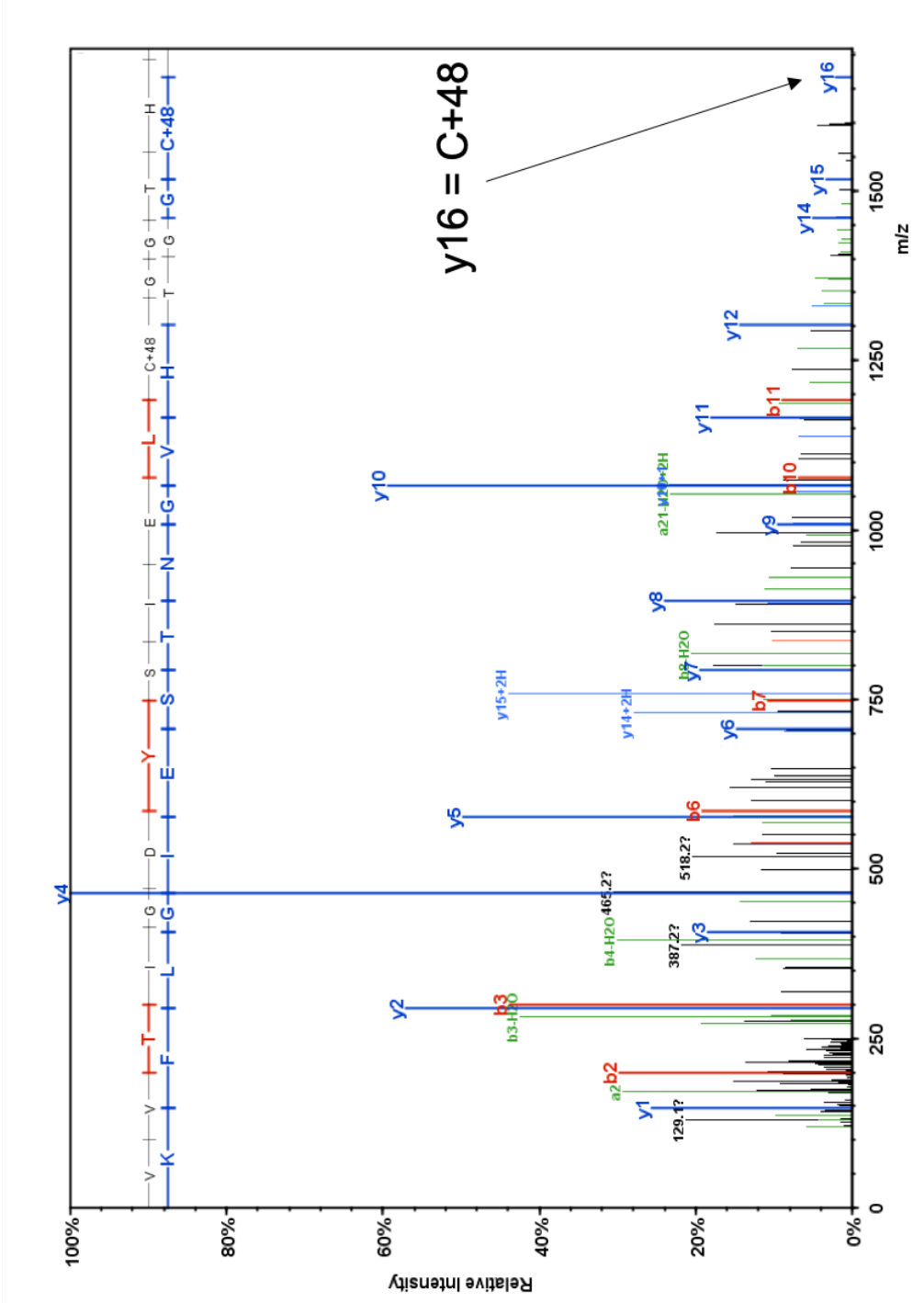



**Figure 7.6: Mass spectra of peptide containing Cys669 from AlaRS control sample.** y16 ion corresponds to Cys+48 Da (indicating tri-oxidation). Data analysed using MaxQuant and Scaffold (Version 4.8.1)

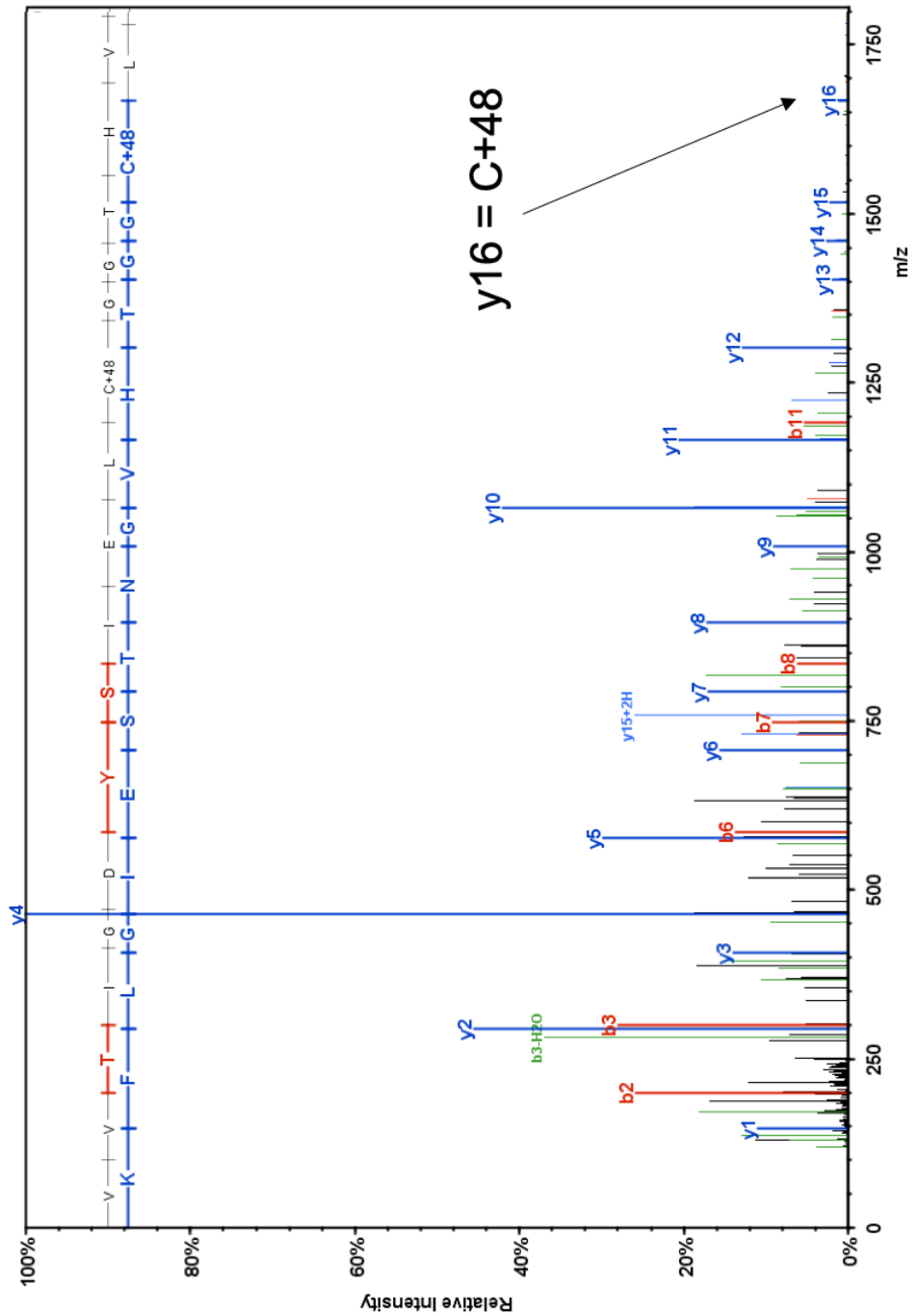



**Figure 7.7: Mass spectra of peptide containing Cys669 from AlaRS $H_2O_2$ treated sample.** y16 ion corresponds to Cys+48 Da (indicating tri-oxidation). Data analysed using MaxQuant and Scaffold (Version 4.8.1)

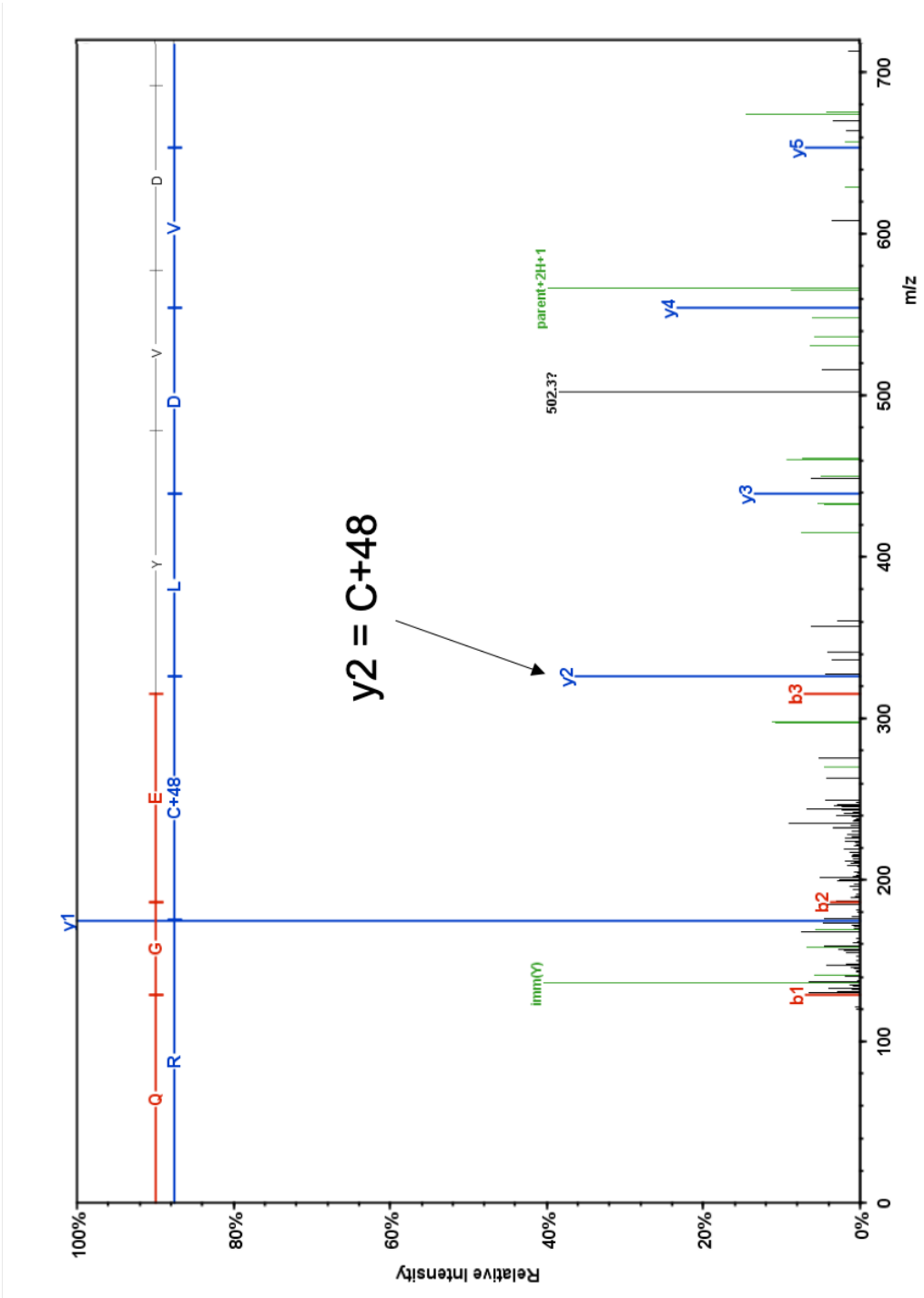



**Figure 7.8: Mass spectra of peptide containing Cys181 from ThrRS control sample.** y2 ion corresponds to Cys+48 Da (indicating tri-oxidation). Data analysed using MaxQuant and Scaffold (Version 4.8.1)

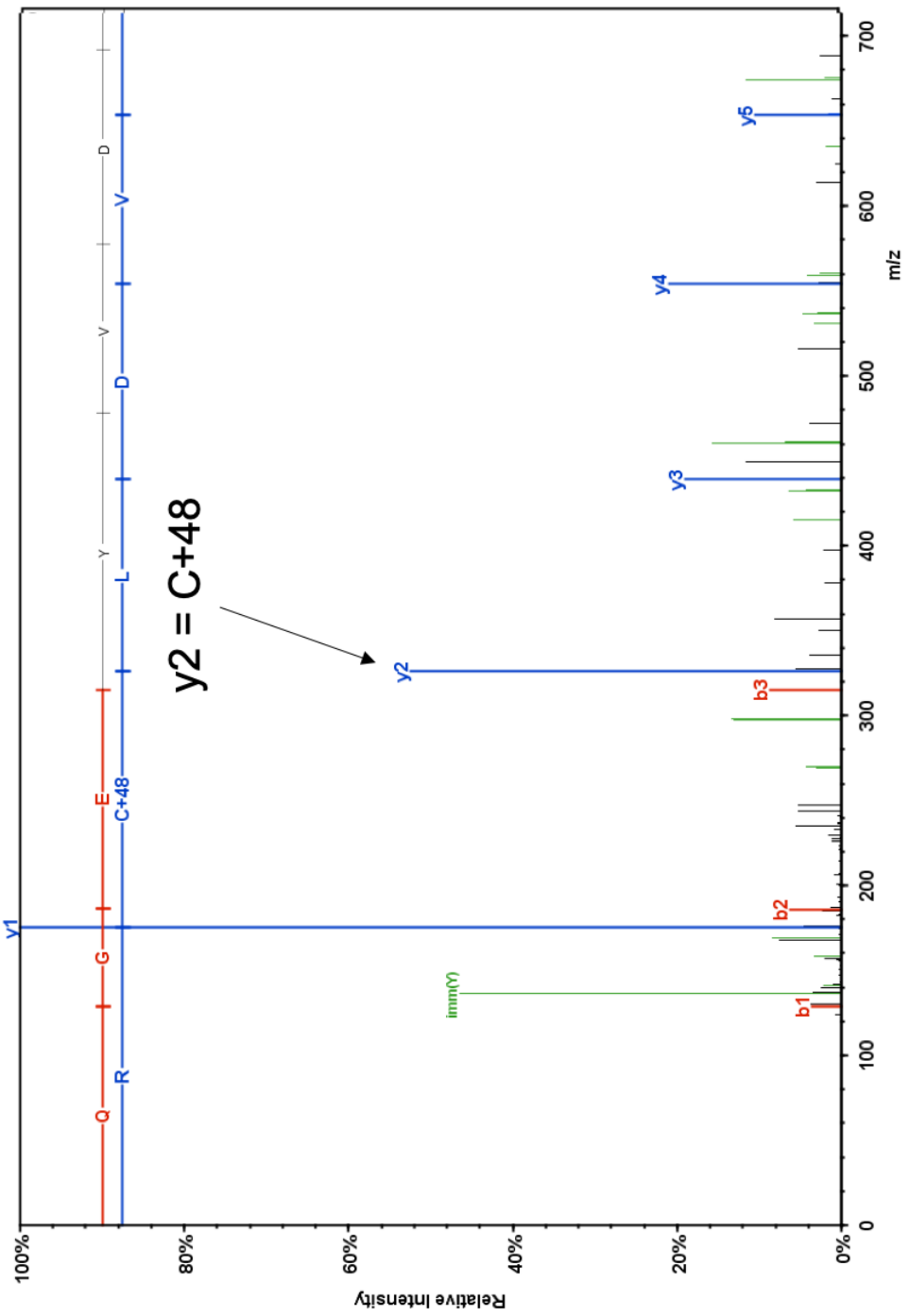



**Figure 7.9: Mass spectra of peptide containing Cys181 from ThrRS $H_2O_2$ treated sample.** y2 ion corresponds to Cys+48 Da (indicating tri-oxidation). Data analysed using MaxQuant and Scaffold (Version*)

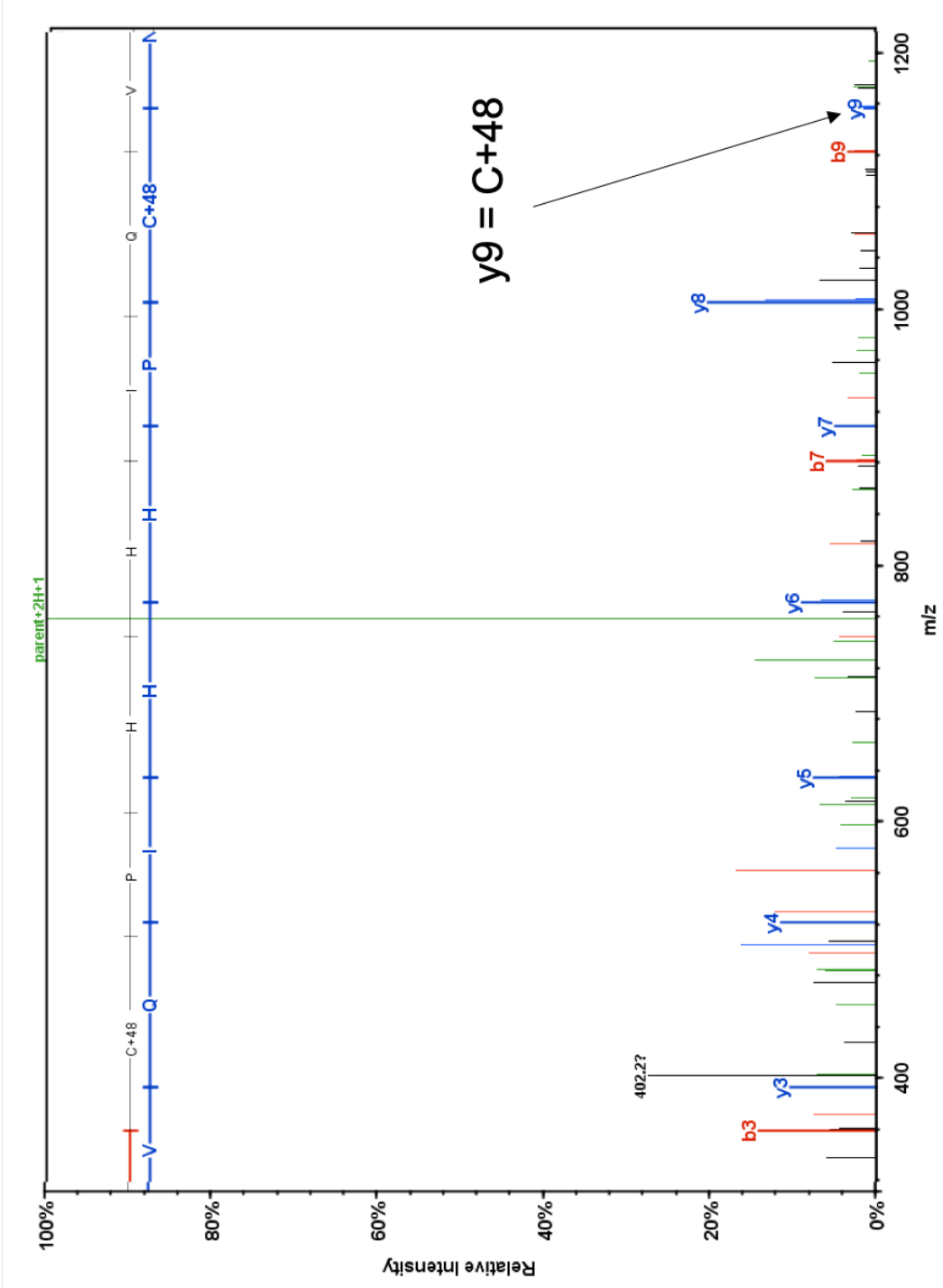



**Figure 7.10: Mass spectra of peptide containing Cys336 from ThrRS control sample.** y9 ion corresponds to Cys+48 Da (indicating tri-oxidation). Data analysed using MaxQuant and Scaffold (Version 4.8.1)

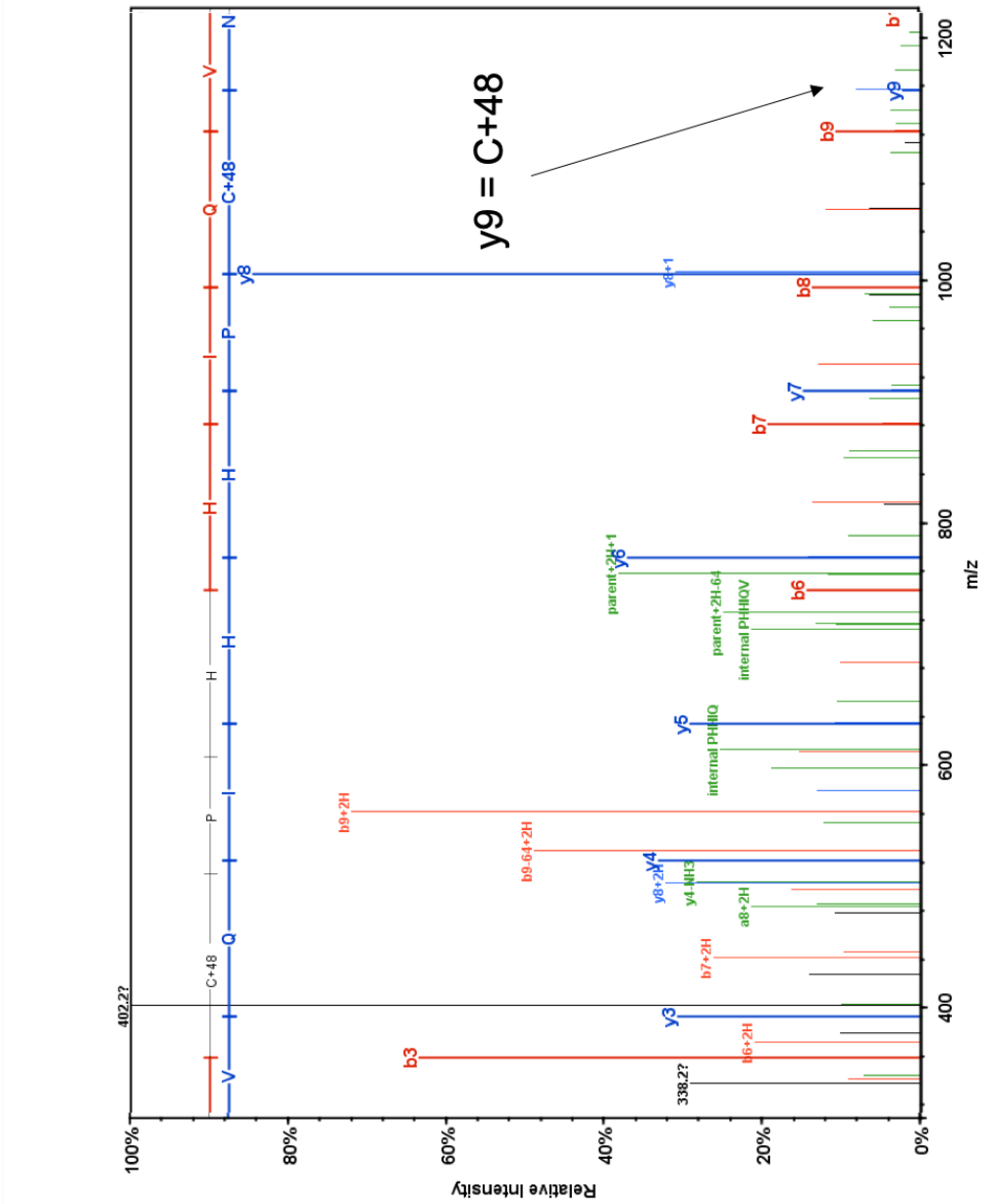



**Figure 7.11: Mass spectra of peptide containing Cys336 from ThrRS $H_2O_2$ treated sample.** y9 ion corresponds to Cys+48 Da (indicating tri-oxidation). Data analysed using MaxQuant and Scaffold (Version 4.8.1)

## 7.5 Structures of synthesised substrates

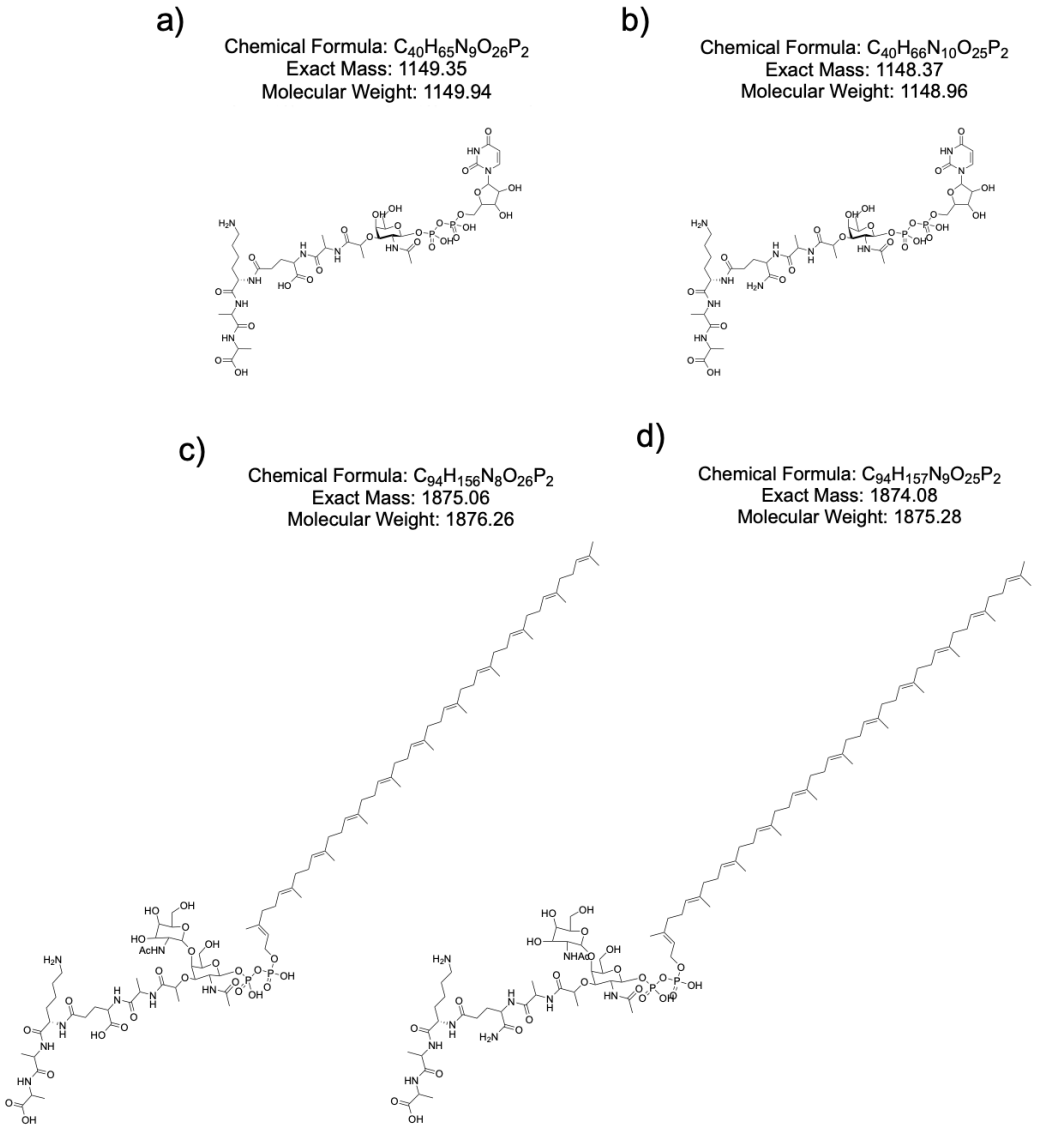



**Figure 7.12: Chemical structures of synthesised substrates.** a) UDP-*N*-acetylmuramyl pentapeptide (UDP-5P), b) iGln UDP-*N*-acetylmuramyl pentapeptide (UDP-5P), c) Lipid II(lys) and d) Lipid II(lys). Figure created using ChemDraw Professional (Version 17.1.1.0).

## 7.6 tRNA gene alignments in *S. pneumoniae*

| tRNA | Length (bp) | Extinction Coefficient | Molecular Weight (Da) |
|---|---|---|---|
| $tRNA^{Ala}$ | 76 | 982.6 | 23,578 |
| $tRNA^{Ser1}$ | 91 | 1179.2 | 28,409 |
| $tRNA^{Ser2}$ | 91 | 1178.5 | 28,352 |
| $tRNA^{Ser3}$ | 93 | 1216.9 | 28,872 |
| $tRNA^{Thr1}$ | 76 | 985 | 23,607 |
| $tRNA^{Thr2}$ | 76 | 971.1 | 23,644 |

**Table 7.3: Molecular features of tRNA isoacceptors.** The isoacceptor, length, extinction coefficient and molecular weight are shown.

## 7.7 Structure of truncated Lipid II(lys)

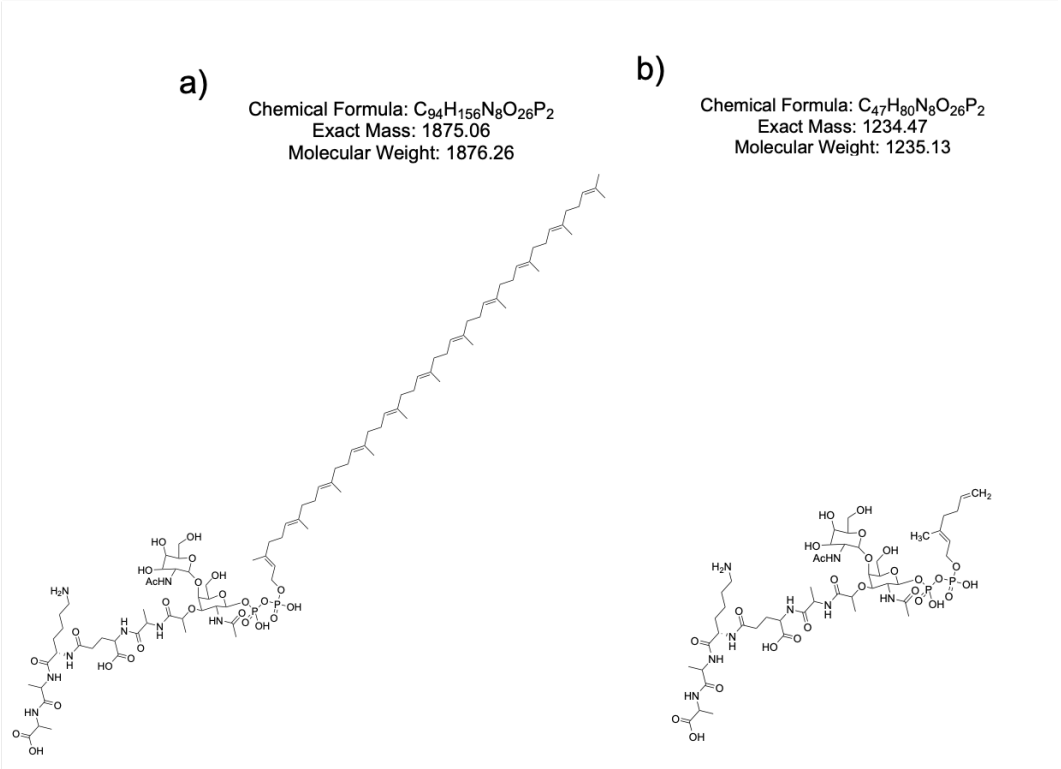



**Figure 7.13: Structure of MurM substrate and truncated version for molecular docking.** a) Lipid II(lys) - the true substrate of MurM, b) Truncated Lipid II(lys) - containing a short hydrocarbon tail required for molecular docking predictions. Figure created using ChemDraw Professional (Version 17.1.1.0).

## 7.8 Discrete optimized protein energy (DOPE-HR) profile

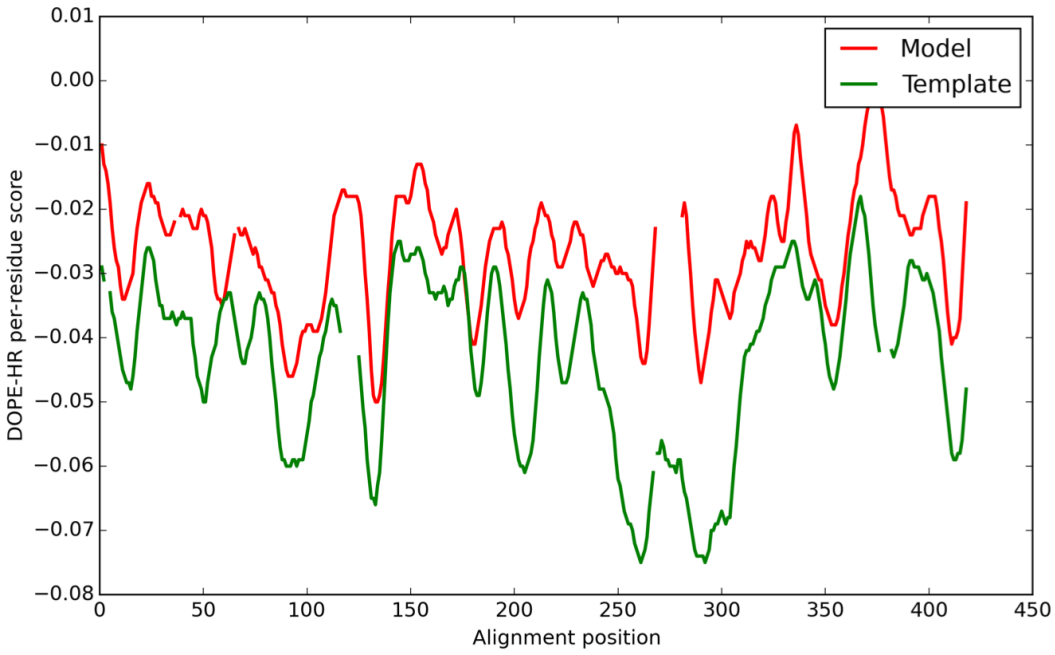



**Figure 7.14: Discrete Optimized Protein Energy Profile for MurM and FemX.** Comparison of DOPE-HR profiles for MurM model (red) and FemX template (green).

## 7.9 Depletion-enrichment (D-E) index for simulated membrane lipids

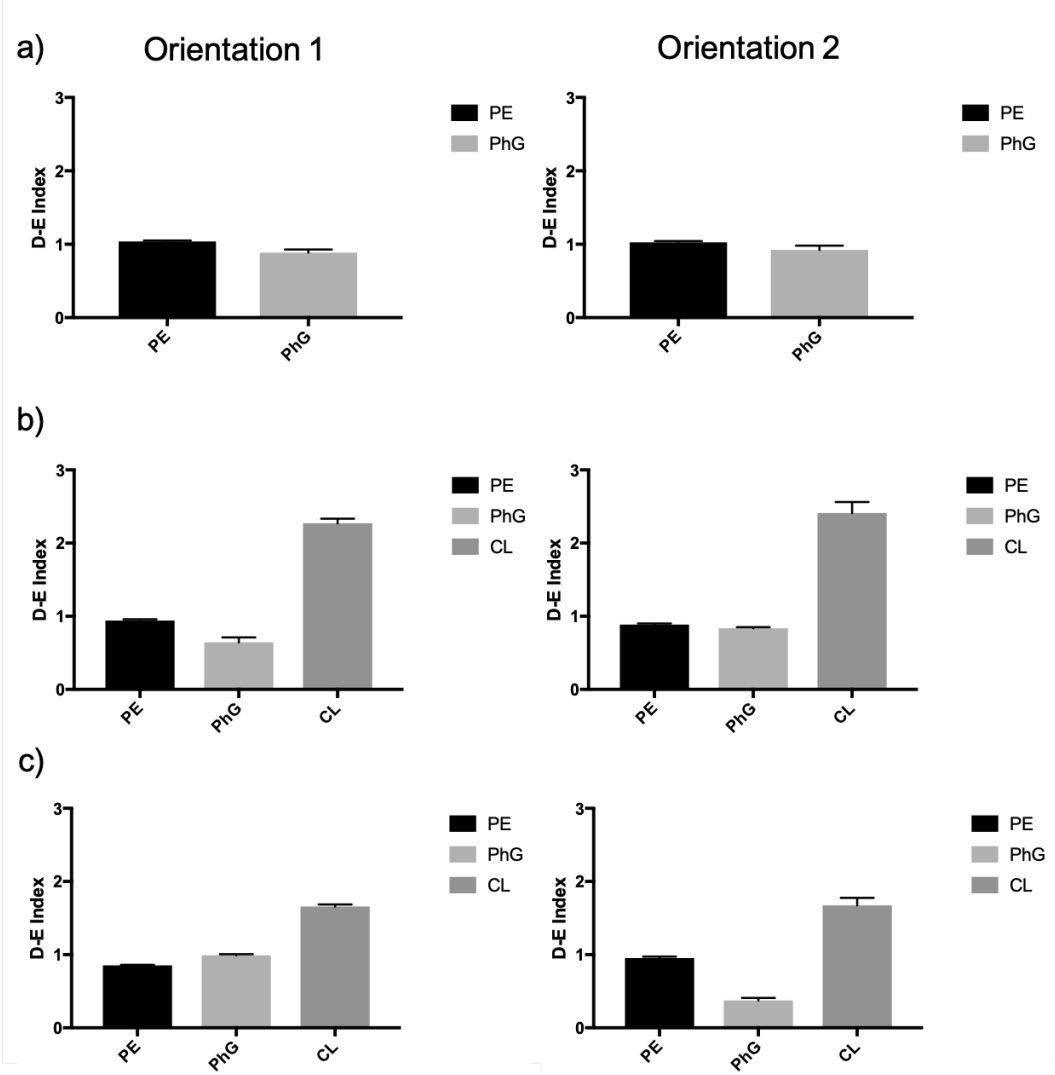



**Figure 7.15: D-E index full data set.** D-E index for phosphatidylethanolamine (POPE), phosphotidylglycerol (POPG) and cardiolipin (CL) with MurM simulations beginning in two alternate starting orientations for a) system 1, b) system 2 and c) system 3.

## 7.10 Starting orientations for MurM in all-atom simulations

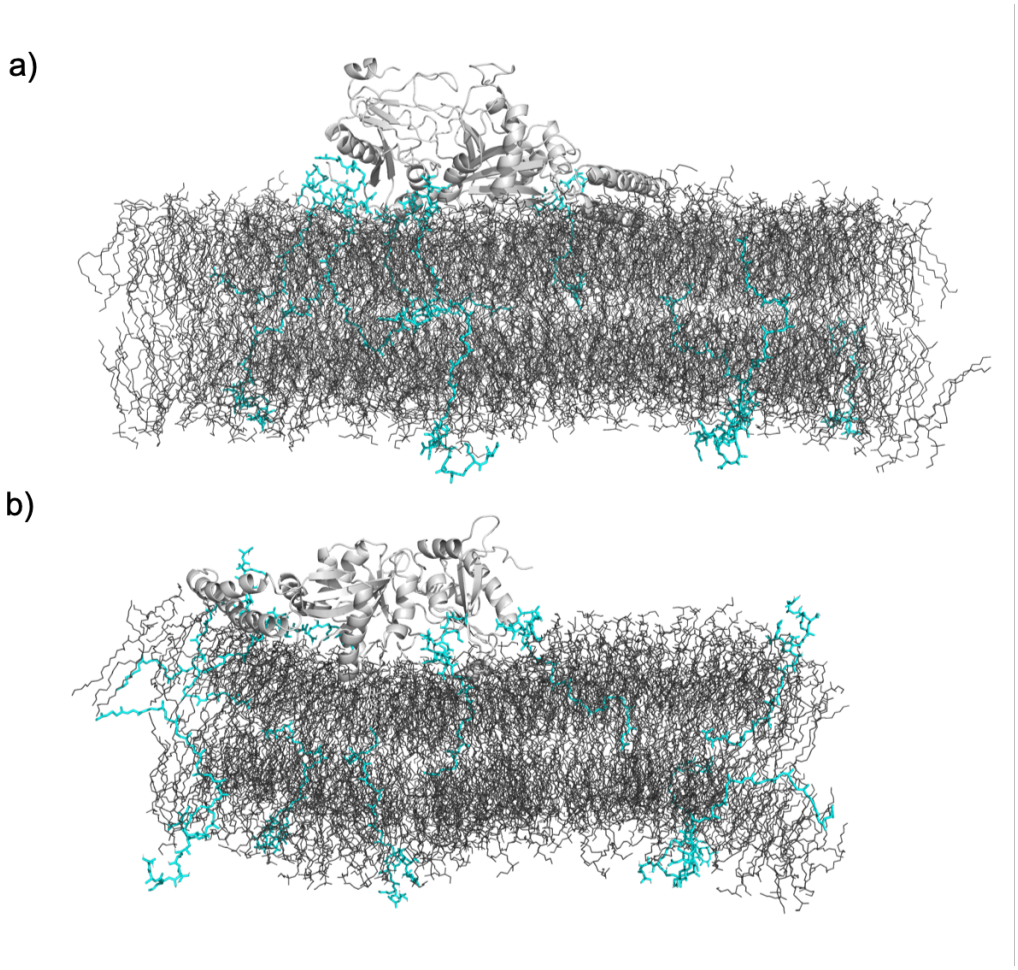



**Figure 7.16: Starting orientations of MurM on the membrane.** Alternate orientations of MurM interactions with the membrane, as determined by coarse-grain simulations. Both orientations used for all-atom simulations a) binding site facing towards the membrane and b) binding site facing away from the membrane. Figure made using PyMOL (Version 2.2.0).

## 7.11 Plasmid maps

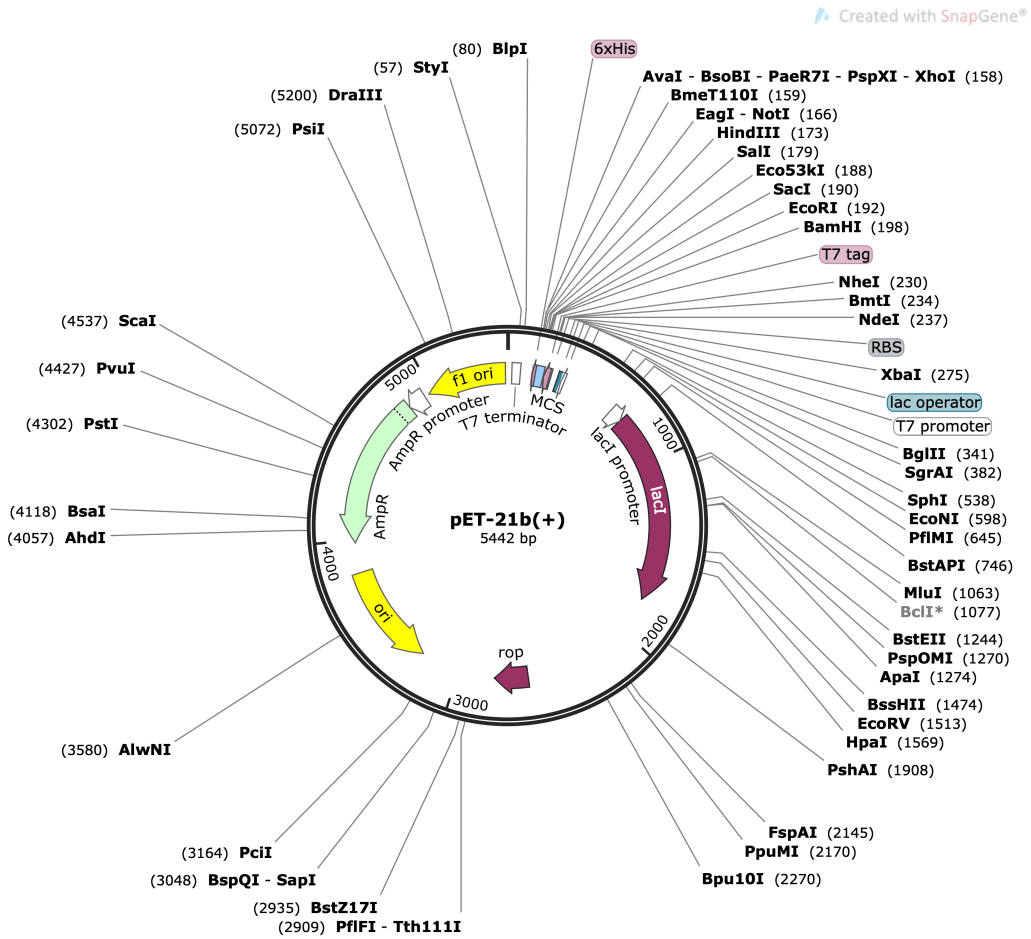



**Figure 7.17: Plasmid map for the bacterial expression vector pET-21b(+).** The origin of replication *(ori)*, ampicillin resistance gene *(AmpR)*, promotor and terminator from T7 phage (T7 promotor/terminator), C-terminal hexa-his tag (6xHis), multiple cloning site (MCS), ribosome binding site (RBS), lactose operator, lactose repressor *(lacI)*, replication regulator protein encoding gene *(rop)* and restriction sites are indicated. Figure made using GeneSnap®.

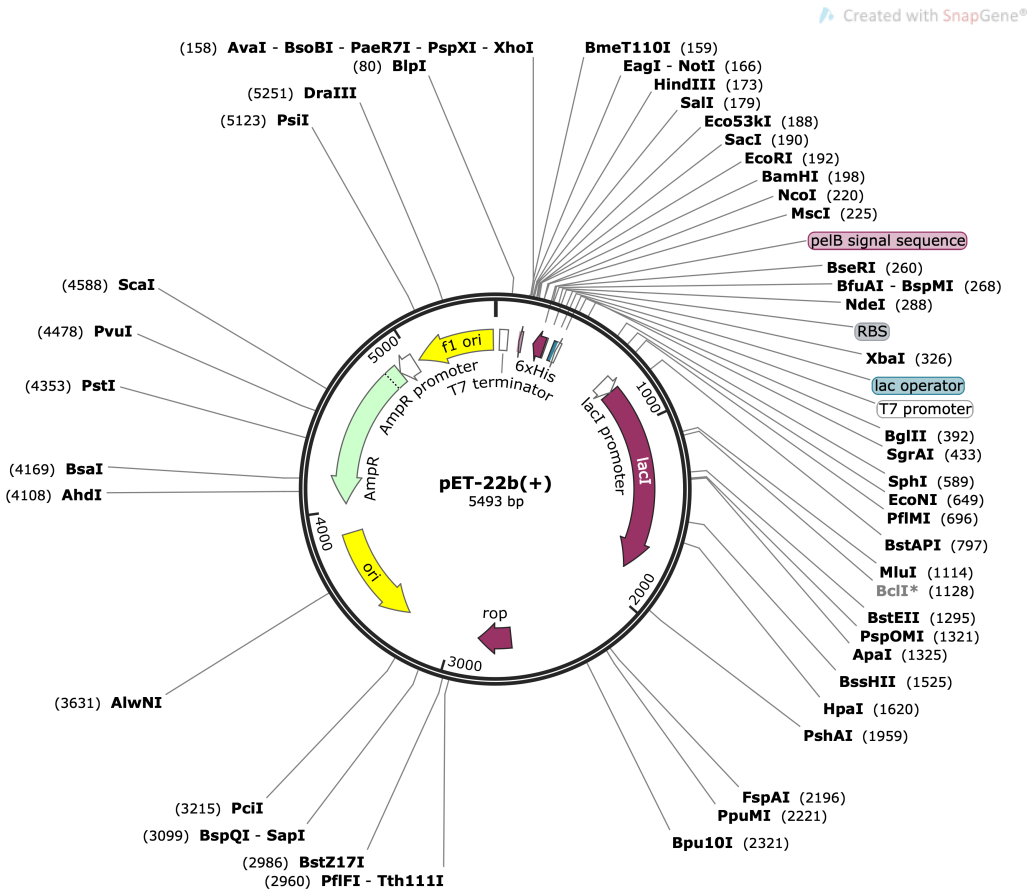



**Figure 7.18: Plasmid map for the bacterial expression vector pET-22b(+).** The origin of replication *(ori)*, ampicillin resistance gene *(AmpR)*, promotor and terminator from T7 phage (T7 promotor/terminator), C-terminal hexa-his tag (6xHis), multiple cloning site (MCS), pelB signal sequence, ribosome binding site (RBS), lactose operator, lactose repressor *(lacI)*, replication regulator protein encoding gene *(rop)* and restriction sites are indicated. Figure made using GeneSnap®.

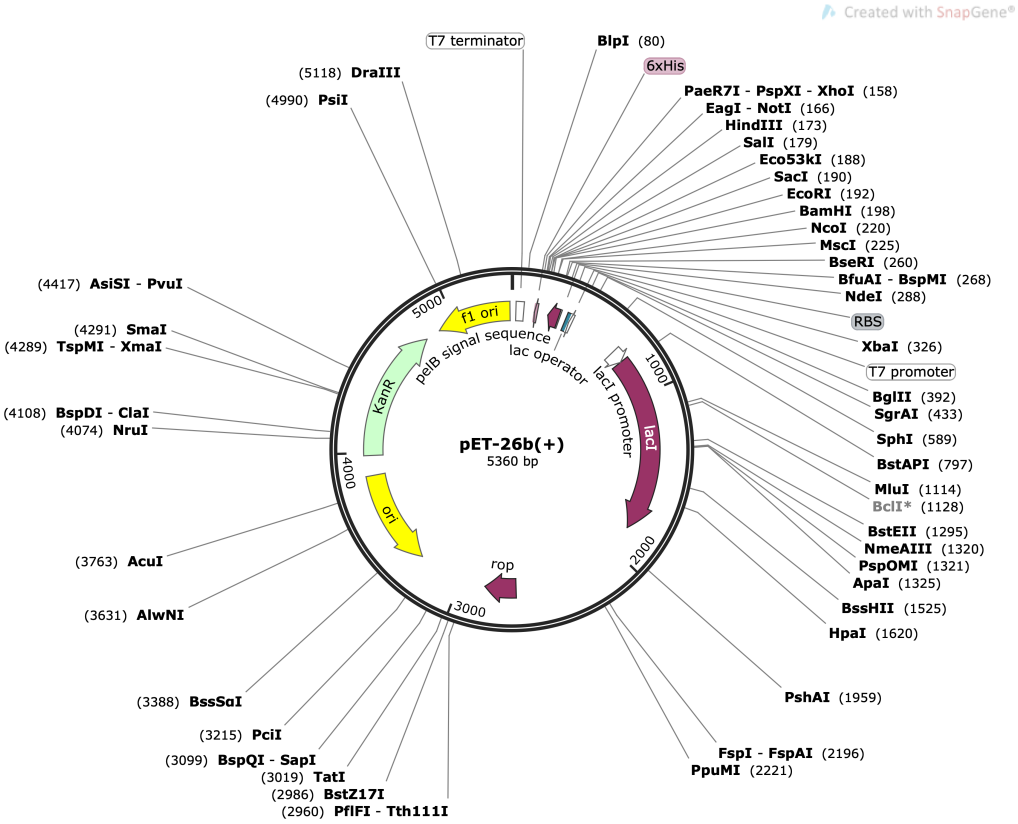



**Figure 7.19: Plasmid map for the bacterial expression vector pET-26b(+).** The origin of replication *(ori)*, kanamycin resistance gene *(KanR)*, promotor and terminator from T7 phage (T7 promotor/terminator), C-terminal hexa-his tag (6xHis), multiple cloning site (MCS), ribosome binding site (RBS), lactose operator, lactose repressor *(lacI)*, replication regulator protein encoding gene *(rop)* and restriction sites are indicated. Figure made using GeneSnap®.

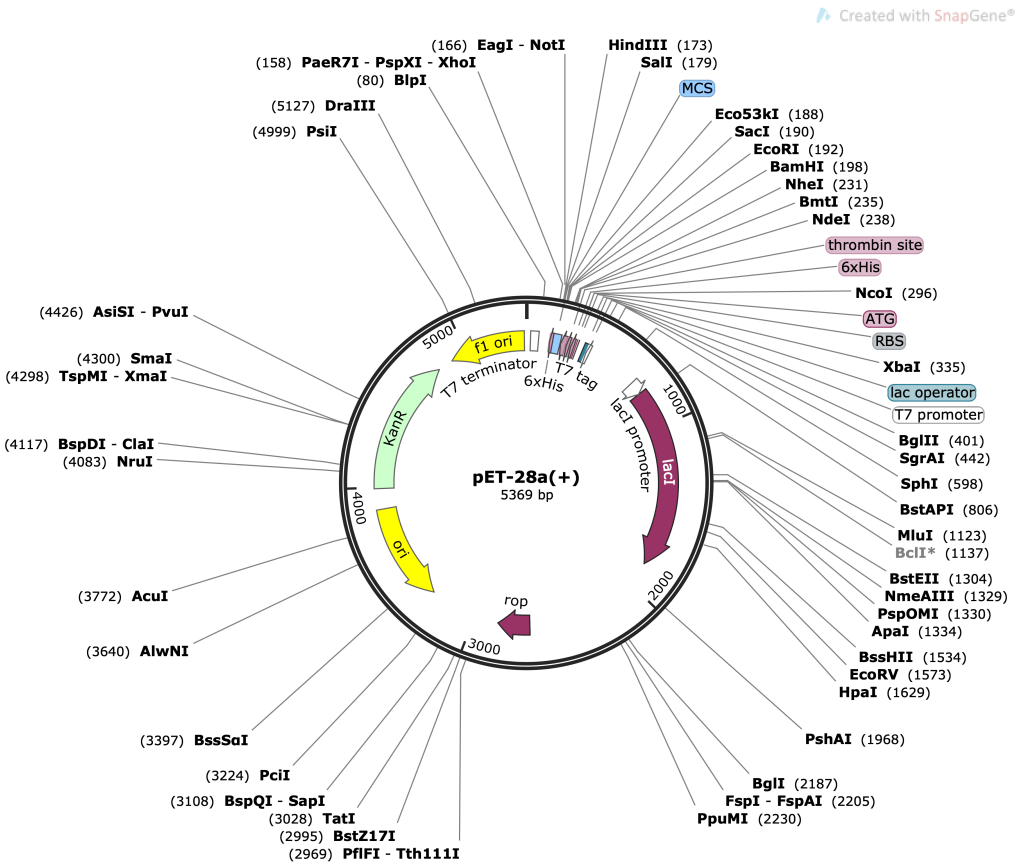



**Figure 7.20: Plasmid map for the bacterial expression vector pET-28a(+).** The origin of replication *(ori)*, kanamycin resistance gene *(KanR)*, promotor and terminator from T7 phage (T7 promotor/terminator), thrombin cleavage site, N- and C-terminal hexa-his tag (6xHis), multiple cloning site (MCS), ribosome binding site (RBS), lactose operator, lactose repressor *(lacI)*, replication regulator protein encoding gene *(rop)* and restriction sites are indicated. Figure made using GeneSnap®.

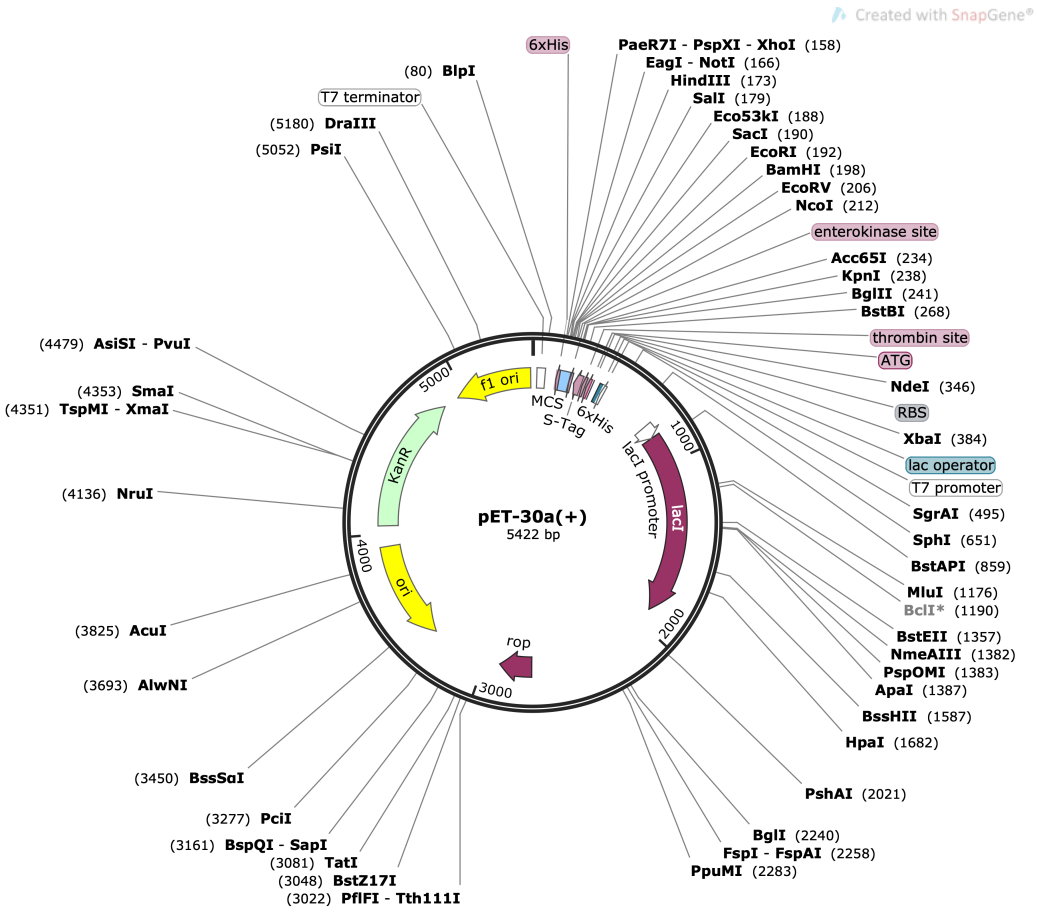



**Figure 7.21: Plasmid map for the bacterial expression vector pET-30a(+).** The origin of replication *(ori)*, kanamycin resistance gene *(KanR)*, promotor and terminator from T7 phage (T7 promotor/terminator), thrombin cleavage site, N- and C-terminal hexa-his tag (6xHis), multiple cloning site (MCS), ribosome binding site (RBS), lactose operator, lactose repressor *(lacI)*, replication regulator protein encoding gene *(rop)*, and restriction sites are indicated. Figure made using GeneSnap®.

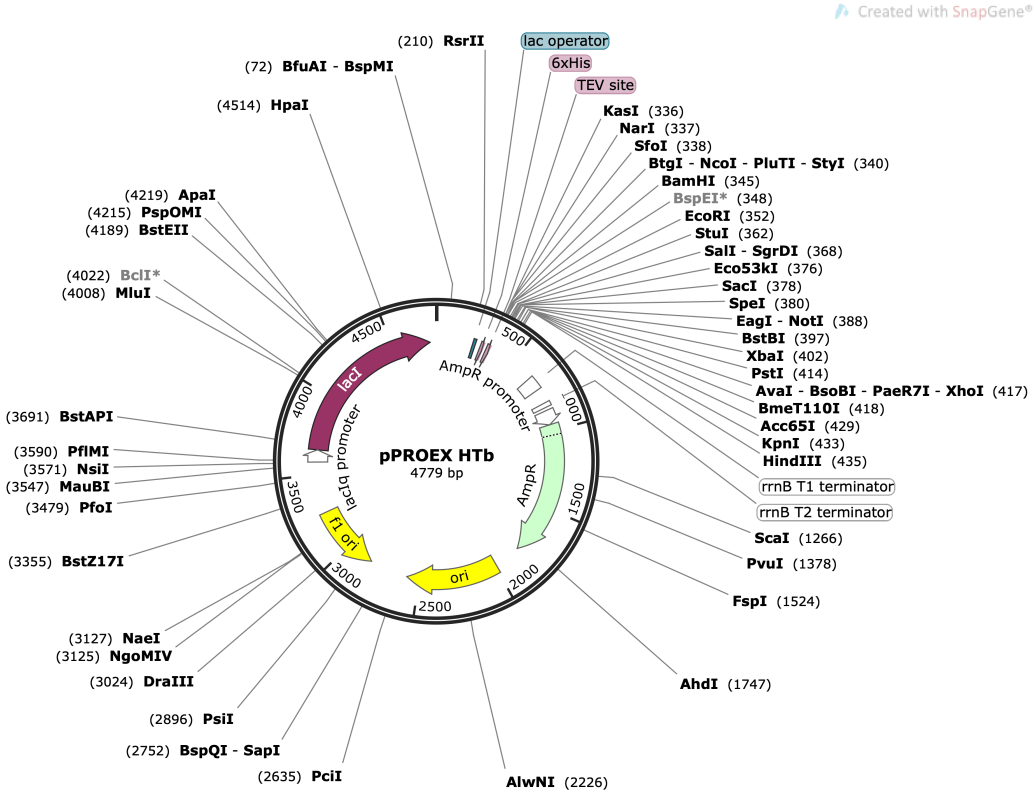



**Figure 7.22: Plasmid map for the bacterial expression vector pET-30a(+).** The origin of replication *(ori)*, lactose repressor *(lacI)*, lactose operator, N-terminal hexa-his tag (6xHis), Tobacco Etch Virus (TEV) cleavage site, promotor and rrnB T1 and T2 terminator, ampicillin resistance gene *(AmpR)*, and restriction sites are indicated. Figure made using GeneSnap®.

# Bibliography

Killing niche competitors by remote-control bacteriophage induction, author=Selva, Laura and Viana, David and Regev-Yochay, Gili and Trzcinski, Krzysztof and Corpa, Juan Manuel and Novick, Richard P and Penadés, José R and others. *Proceedings of the National Academy of Sciences*, 106(4):1234–1238, 2009.

Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015.

Ravi K Acharya and Matthew D Lloyd. The advantages and limitations of protein crystal structures. *Trends in pharmacological sciences*, 26(1):10–14, 2005.

WA Adedeji. The treasure called antibiotics. *Annals of Ibadan postgraduate medicine*, 14(2):56, 2016.

Ivan Ahel, Dragana Korencic, Michael Ibba, and Dieter Söll. Trans-editing of mischarged trnas. *Proceedings of the National Academy of Sciences*, 100(26): 15422–15427, 2003. ISSN 0027-8424. doi: 10.1073/pnas.2136934100. URL `https://www.pnas.org/content/100/26/15422`.

Saheem Ahmad, Hamda Khan, Uzma Shahab, Shahnawaz Rehman, Zeeshan Rafi, Mohd Yasir Khan, Ahsanullah Ansari, Zeba Siddiqui, Jalaluddin Mohammad Ashraf, SM Abdullah, *et al.* Protein oxidation: an overview of metabolism of sulphur containing amino acid, cysteine. *Front. Biosci*, 9:71–87, 2017.

Rustam I Aminov. A brief history of the antibiotic era: lessons learned and challenges for the future. *Frontiers in microbiology*, 1:134, 2010.

Songon An and Karin Musier-Forsyth. Cys-tRNAPro editing by Haemophilus influenzae YbaK via a novel Synthetase· YbaK· tRNA ternary complex. *Journal of Biological Chemistry*, 280(41):34465–34472, 2005.

Konstantin Arnold, Lorenza Bordoli, Jrgen Kopp, and Torsten Schwede. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2):195–201, 2006.

Panu Artimo, Manohar Jonnalagedda, Konstantin Arnold, Delphine Baratin, Gabor Csardi, Edouard De Castro, Séverine Duvaud, Volker Flegel, Arnaud Fortier, Elisabeth Gasteiger, *et al.* ExPASy: SIB bioinformatics resource portal. *Nucleic acids research*, 40(W1):W597–W603, 2012.

Rajat Banerjee, Shawn Chen, Kiley Dare, Marla Gilreath, Mette Praetorius-Ibba, Medha Raina, Noah M. Reynolds, Theresa Rogers, Herv Roy, Srujana S. Yadavalli, and Michael Ibba. tRNAs: Cellular barcodes for amino acids. *FEBS Letters*, 584(2), 2010. ISSN 1873-3468.

Anton Bankevich, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, Sergey I. Nikolenko, Son Pham, Andrey D. Prjibelski, Alexey V. Pyshkin, Alexander V. Sirotkin, Nikolay Vyahhi, Glenn Tesler, Max A. Alekseyev, and Pavel A. Pevzner. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5):455–477, 2012.

Victoria A Barcus, Kiran Ghanekar, Maggie Yeo, Tracey J Coffey, and Christopher G Dowson. Genetics of high level penicillin resistance in clinical isolates of Streptococcus pneumoniae. *FEMS Microbiology Letters*, 126(3):299–303, 1995.

Héléne Barreteau, Andreja Kovac, Audrey Boniface, Matej Sova, Stanislav Gobec,

and Didier Blanot. Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiology Reviews*, 32(2):168–207, 2008.

Kirk Beebe, Lluis Ribas de Pouplana, and Paul Schimmel. Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability. *The EMBO Journal*, 22(3):668–675, 2003.

Samuel Bellais, Michel Arthur, Lionnel Dubost, Jean-Emmanuel Hugonnet, Laurent Gutmann, Jean Van Heijenoort, Raymond Legrand, Jean-Paul Brouard, Louis Rice, and Jean-Luc Mainardi. Aslfm, the D-aspartate ligase responsible for the addition of D-aspartic acid onto the peptidoglycan precursor of Enterococcus faecium. *Journal of Biological Chemistry*, 281(17):11586–11594, 2006.

Timothy E Benson, D Bryan Prince, Veronica T Mutchler, Kimberly A Curry, Andrea M Ho, Ronald W Sarver, Jeanne C Hagadorn, Gil H Choi, and Robert L Garlick. X-ray crystal structure of Staphylococcus aureus FemA. *Structure*, 10(8): 1107–1115, 2002.

Herman JC Berendsen, JPM van Postma, Wilfred F van Gunsteren, ARHJ DiNola, and JR Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.

Jeremy M Berg, John L Tymoczko, and Lubert Stryer. *Biochemistry*. New York: W H Freeman, 5 edition, 2002. ISBN 0-7167-3051-0.

Brigitte Berger-Bächi and Martin Tschierske. Role of fem factors in methicillin resistance. *Drug Resistance Updates*, 1(5):325–335, 1998.

Giuseppe Bertani. Studies on lysogenesis.i: The mode of phage liberation by lysogenic Escherichia coli. *Journal of bacteriology*, 62(3):293, 1951.

Penny J Beuning and Karin Musier-Forsyth. Hydrolytic editing by a class II aminoacyl-tRNA synthetase. *Proceedings of the National Academy of Sciences*, 97(16): 8916–8920, 2000.

Sabrina Biarrotte-Sorin, Antoine P Maillard, Jean Delettré, Wladimir Sougakoff,

Michel Arthur, and Claudine Mayer. Crystal structures of Weissella viridescens FemX and its complex with UDP-MurNAc-pentapeptide: insights into FemABX family substrates recognition. *Structure*, 12(2):257–267, 2004.

Valrie Biou, Anna Yaremchuk, Michael Tukalo, and Stephen Cusack. The 2.9 crystal structure of T. thermophilus Seryl-tRNA synthetase complexed with $tRNA^{Ser}$. *Science*, 263(5152):1404–1410, 1994.

Glenn R Björk, Johanna U Ericson, Claes ED Gustafsson, Tord G Hagervall, Yvonne H Jönsson, and P Mikael Wikström. Transfer RNA modification. *Annual review of biochemistry*, 56(1):263–285, 1987.

Jessica MA Blair, Mark A Webber, Alison J Baylay, David O Ogbolu, and Laura JV Piddock. Molecular mechanisms of antibiotic resistance. *Nature reviews microbiology*, 13(1):42, 2015.

Jan-Willem P Boots, Robin EJ Spelbrink, Gerda M Kool, Eefjan Breukink, J Antoinette Killian, Ben de Kruijff, *et al.* Membrane interaction of the glycosyltransferase MurG: a special role for cardiolipin. *Journal of bacteriology*, 185(13):3773–3779, 2003.

Franck Borel, Christine Vincent, Reuben Leberman, and Michael Härtlein. Seryl-tRNA synthetase from Escherichia coli: implication of its N-terminal domain in aminoacylation activity and specificity. *Nucleic acids research*, 22(15):2963–2969, 1994.

F Bouadloun, D Donner, and CG Kurland. Codon-specific missense errors in vivo. *The EMBO journal*, 2(8):1351–1356, 1983.

Ahmed Bouhss, Nathalie Josseaume, Anatoly Severin, Keiko Tabei, Jean-Emmanuel Hugonnet, David Shlaes, Dominique Mengin-Lecreulx, Jean Van Heijenoort, and Michel Arthur. Synthesis of the L-alanyl-L-alanine cross-bridge of Enterococcus faecalis peptidoglycan. *J Biol Chem*, 277(48):4593545941, 2002.

George EP Box and Norman R Draper. *Empirical model-building and response surfaces.* John Wiley & Sons, 1987.

Eefjan Breukink, Hester E van Heusden, Pauline J Vollmerhaus, Ewa Swiezewska, Livia Brunne, Suzanne Walke, Albert J. R Heck, and Ben de Kruijff. Lipid II Is an Intrinsic Component of the Pore Induced by Nisin in Bacterial Membranes. *Journal of Biological Chemistry*, 278(22):19898–903, 2003.

Reinhold Brückner, Michael Nuhn, Peter Reichmann, Beate Weber, and Regine Hakenbeck. Mosaic genes and mosaic chromosomesgenomic variation in streptococcus pneumoniae. *International Journal of Medical Microbiology*, 294(2):157 – 168, 2004. ISSN 1438-4221. doi: https://doi.org/10.1016/j.ijmm.2004.06.019. URL `http://www.sciencedirect.com/science/article/pii/S1438422104000505`.

Nienke Buddelmeijer and Jon Beckwith. A complex of the Escherichia coli cell division proteins FtsL, FtsB and FtsQ forms independently of its localization to the septal region. *Molecular microbiology*, 52(5):1315–1327, 2004.

TD Bugg, D Braddick, CG Dowson, and DI Roper. Bacterial cell wall assembly: still an attractive antibacterial target. *Trends Biotechnol.*, 29(4):167–73, 2011.

Tammy J Bullwinkle, Noah M Reynolds, Medha Raina, Adil Moghal, Eleftheria Matsa, Andrei Rajkovic, Huseyin Kayadibi, Farbod Fazlollahi, Christopher Ryan, Nathaniel Howitz, *et al.* Oxidation of cellular amino acid pools leads to cytotoxic mistranslation of the genetic code. *Elife*, 3:e02501, 2014.

Robert M Bumsted, June L Dahl, Dieter Söll, and Jack L Strominger. Biosynthesis of the peptidoglycan of bacterial cell walls X. Further study of the glycyl transfer ribonucleic acids active in peptidoglycan synthesis in Staphylococcus aureus. *Journal of Biological Chemistry*, 243(4):779–782, 1968.

Jonathan Burbaum, Ruth Starzyk, and Paul Schimmel. Understanding structural relationships proteins of unsolved three-dimensional structure. *Proteins*, 7:99–111, 01 1990.

Centers for Disease Control and Prevention Contributors. Active Bacterial Core Surveillence, 2019. URL `https://wwwn.cdc.gov/BactFacts/index.html`. [Online; accessed 25-March-2019].

Patricia P Chan and Todd M Lowe. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic acids research*, 37:D93–D97, 2009.

Vincent B Chen, W Bryan Arendall, Jeffrey J Headd, Daniel A Keedy, Robert M Immormino, Gary J Kapral, Laura W Murray, Jane S Richardson, and David C Richardson. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66 (1):12–21, 2010.

Carol E Chenoweth, Sanjay Saint, Fernando Martinez, Joseph P Lynch III, and A Mark Fendrick. Antimicrobial resistance in Streptococcus pneumoniae: implications for patients with community-acquired pneumonia. 75(11):1161–1168, 2000.

Sarah M. Chiang and Herb E. Schellhorn. Regulators of oxidative stress response genes in Escherichia coli and their functional conservation in bacteria. *Archives of Biochemistry and Biophysics*, 525(2):161 – 169, 2012.

CLSI.

Valentina Corradi, Eduardo Mendez-Villuendas, Helgi I Inglfsson, Ruo-Xu Gu, Iwona Siuda, Manuel N Melo, Anastassiia Moussatova, Lucien J DeGagn, Besian I Sejdiu, Gurpreet Singh, *et al.* Lipid–protein interactions are unique fingerprints for membrane proteins. *ACS central science*, 4(6):709–717, 2018.

Vitor Costa, Alexandre Quintanilha, and Pedro Moradas-Ferreira. Protein oxidation, repair mechanisms and proteolysis in Saccharomyces cerevisiae. *IUBMB life*, 59 (4-5):293–298, 2007.

FHC Crick. Codon-anticodon pairing: the wobble hypothesis. 1966.

Yadong A Cui, Harshila Patel, William M. O'Neil, Se Li, and Patricia Saddier.

Pneumococcal serotype distribution: A snapshot of recent data in pediatric and adult populations around the world. *Human vaccines immunotherapeutics*, 13(6): 1229–1241, 2017.

Djurre H de Jong, Gurpreet Singh, WF Drew Bennett, Clement Arnarez, Tsjerk A Wassenaar, Lars V Schfer, Xavier Periole, D Peter Tieleman, and Siewert J Marrink. Improved parameters for the martini coarse-grained protein force field. *Journal of Chemical Theory and Computation*, 9(1):687–697, 2012.

Anne-Catherine Dock-Bregeon, Rajan Sankaranarayanan, Pascale Romby, Joel Caillet, Mathias Springer, Bernard Rees, Christopher S Francklyn, Chantal Ehresmann, and Dino Moras. Transfer RNA-mediated editing in threonyl-tRNA synthetase: the class II solution to the double discrimination problem. *Cell*, 103(6): 877–884, 2000.

Guang Qiang Dong, Hao Fan, Dina Schneidman-Duhovny, Ben Webb, and Andrej Sali. Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics*, 29(24):3158–3166, 2013.

Christopher G Dowson, Kiran Ghanekar, Maggie Yeo, Tracey J Coffey, and Victoria A Barcus. Genetics of high level penicillin resistance in clinical isolates of Streptococcus pneumoniae. *FEMS Microbiology Letters*, 126(3):299–303, 03 1995. ISSN 0378-1097. doi: 10.1111/j.1574-6968.1995.tb07433.x. URL `https://doi.org/10.1111/j.1574-6968.1995.tb07433.x`.

Shaynoor Dramsi, Sophie Magnet, Sophie Davison, and Michel Arthur. Covalent attachment of proteins to peptidoglycan. *FEMS Microbiology Reviews*, 32(2): 307–320, 02 2008. ISSN 0168-6445. doi: 10.1111/j.1574-6976.2008.00102.x. URL `https://doi.org/10.1111/j.1574-6976.2008.00102.x`.

Morana Dulic, Nevena Cvetesic, John J Perona, and Ita Gruic-Sovulj. Partitioning of tRNA-dependent editing between pre-and post-transfer pathways in class I aminoacyl-trna synthetases. *Journal of Biological Chemistry*, 285(31):23799–23809, 2010.

Patricia Edelmann and Jonathan Gallant. Mistranslation in e. coli. *Cell*, 10(1):131–137, 1977.

K. Ehlert, W. Schroder, and H. Labischinski. Specificities of FemA and FemB for different glycine residues: FemB cannot substitute for FemA in staphylococcal peptidoglycan pentaglycine side chain formation. *Journal of Bacteriology*, 179(23): 7573–7576, 1997. ISSN 00219193. doi: 10.1128/jb.179.23.7573-7576.1997.

Ahmed El Zoeiby, Francois Sanschagrin, Josëe Lamoureux, Andrë Darveau, and Roger C Levesque. Cloning, over-expression and purification of pseudomonas aeruginosa murc encoding uridine diphosphate n-acetylmuramate: L-alanine ligase. *FEMS microbiology letters*, 183(2):281–288, 2000.

Ahmed El Zoeiby, François Sanschagrin, Pierre C Havugimana, Alain Garnier, and Roger C Levesque. In vitro reconstruction of the biosynthetic pathway of peptidoglycan cytoplasmic precursor in pseudomonas aeruginosa. *FEMS microbiology letters*, 201(2):229–235, 2001.

Markus Englert, Sarath Moses, Michael Hohn, Jiqiang Ling, Patrick ODonoghue, and Dieter Söll. Aminoacylation of tRNA 2-or 3-hydroxyl by phosphoseryl-and pyrrolysyl-tRNA synthetases. *FEBS letters*, 587(20):3360–3364, 2013.

Gilbert Eriani, Marc Delarue, Olivier Poch, Jean Gangloff, and Dino Moras. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature*, 347(6289):203, 1990.

Narayanan Eswar, Ben Webb, Marc A Marti-Renom, MS Madhusudhan, David Eramian, Min-yi Shen, Ursula Pieper, and Andrej Sali. Comparative protein structure modeling using modeller. *Current protocols in bioinformatics*, 15(1):5–6, 2006.

EUCAST.

Joseph O Falkinham, Thomas E Wall, Justin R Tanner, Khaled Tawaha, Feras Q Alali, Chen Li, and Nicholas H Oberlies. Proliferation of antibiotic-producing bacteria

and concomitant antibiotic production as the basis for the antibiotic activity of jordan's red soils. *Appl. Environ. Microbiol.*, 75(9):2735–2741, 2009.

Thalia A Farazi, Gabriel Waksman, and Jeffrey I Gordon. Structures of saccharomyces cerevisiae N-myristoyltransferase with bound myristoyl CoA and peptide provide insights about substrate recognition and catalysis. *Biochemistry*, 40(21):6335–6343, 2001.

Daniel R. Feikin and Keith P. Klugman. Historical Changes in Pneumococcal Serogroup Distribution: Implications for the Era of Pneumococcal Conjugate Vaccines. *Clinical Infectious Diseases*, 35(5):547–555, 09 2002. ISSN 1058-4838. doi: 10.1086/341896. URL `https://doi.org/10.1086/341896`.

Sergio R. Filipe and Alexander Tomasz. Inhibition of the Expression of Penicillin Resistance in Streptococcus Pneumoniae by Inactivation of Cell Wall Muropeptide Branching Genes. *Proceedings of the National Academy of Sciences of the United States of America*, 97(9):48914896, 2000.

Sergio R Filipe, Mariana G Pinho, and Alexander Tomasz. Characterization of the murMN operon involved in the synthesis of branched peptidoglycan peptides in Streptococcus pneumoniae. *Journal of Biological Chemistry*, 275(36):27768–27774, 2000a.

Sergio R Filipe, Elena Severina, and Alexander Tomasz. Distribution of the Mosaic Structured murM Genes among Natural Populations of *Streptococcus pneumoniae*. *Journal of Bacteriology*, 182(23):67986805, 2000b.

Sergio R Filipe, Elena Severina, and A Tomasz. The role of murMN operon in penicillin resistance and antibiotic tolerance of Streptococcus pneumoniae. *Microbial drug resistance*, 7(4):303–316, 2001a.

Sergio R. Filipe, Elena Severina, and Alexander Tomasz. Functional Analysis of Streptococcus pneumoniae MurM Reveals the Region Responsible for Its

Specificity in the Synthesis of Branched Cell Wall Peptides. *Journal of Biological Chemistry*, 276:39618–39628, 2001b.

Sergio R Filipe, Elena Severina, and Alexander Tomasz. The murMN operon: a functional link between antibiotic resistance and antibiotic tolerance in Streptococcus pneumoniae. *Proceedings of the National Academy of Sciences*, 99 (3):1550–1555, 2002.

Marcus Fischer, Brian K Shoichet, and James S Fraser. One crystal, two temperatures: Cryocooling penalties alter ligand binding to transient protein sites. *ChemBioChem*, 16(11):1560–1564, 2015.

András Fiser, Richard Kinh Gian Do, *et al.* Modeling of loops in protein structures. *Protein science*, 9(9):1753–1773, 2000.

Andrs Fiser, Sergio R. Filipe, and Alexander Tomasz. Cell wall branches, penicillin resistance and the secrets of the MurM protein. *Trends in Microbiology*, 11(12):547 – 553, 2003.

Matthieu Fonvielle, Maryline Chemama, Maxime Lecerf, Régis Villet, Patricia Busca, Ahmed Bouhss, Mélanie Ethève-Quelquejeu, and Michel Arthur. Decoding the Logic of the tRNA Regiospecificity of Nonribosomal FemXWv Aminoacyl Transferase. *Angewandte Chemie International Edition*, 49(30):5115–5119, 2010.

Matthieu Fonvielle, Inés Li de La Sierra-Gallay, Afaf H El-Sagheer, Maxime Lecerf, Delphine Patin, Dénia Mellal, Claudine Mayer, Didier Blanot, Nittaya Gale, Tom Brown, *et al.* The Structure of FemXWv in Complex with a Peptidyl-RNA Conjugate: Mechanism of Aminoacyl Transfer from Ala-tRNAAla to Peptidoglycan Precursors. *Angewandte Chemie*, 125(28):7419–7422, 2013.

Christopher S Francklyn. DNA polymerases and aminoacyl-tRNA synthetases: shared mechanisms for ensuring the fidelity of gene expression. *Biochemistry*, 47 (45):11695–11703, 2008.

Christopher S Francklyn, Eric A First, John J Perona, and Ya-Ming Hou. Methods for kinetic and thermodynamic analysis of aminoacyl-tRNA synthetases.

James S Fraser, Henry van den Bedem, Avi J Samelson, P Therese Lang, James M Holton, Nathaniel Echols, and Tom Alber. Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proceedings of the National Academy of Sciences*, 108(39):16247–16252, 2011.

DN Ganchev, HE Hasper, E Breukink, and B de Kruijff. Size and orientation of the lipid II headgroup as revealed by AFM imaging. *Biochemistry*, 45(19):6195–6202, 2006.

Jose Garcia-Bustos and Alexander Tomasz. A Biological Price of Antibiotic Resistance: Major Changes in the Peptidoglycan Structure of Penicillin-Resistant Pneumococci. *Proceedings of the National Academy of Sciences of the United States of America*, 87(14):5415–5419, 1990.

Elisabeth Gasteiger, Christine Hoogland, Alexandre Gattiker, Marc R Wilkins, Ron D Appel, Amos Bairoch, *et al.* Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook*, pages 571–607. Springer, 2005.

Hoogland C. Ivanyi I. Appel R.D. Bairoch A. Gasteiger E., Gattiker A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*, 31:3784–3788, 2003.

K. Aaron Geno, Gwendolyn L. Gilbert, Joon Young Song, Ian C. Skovsted, Keith P. Klugman, Christopher Jones, Helle B. Konradsen, and Moon H. Nahm. Pneumococcal Capsules and Their Types: Past, Present, and Future. *Clinical Microbiology Reviews*, 28(3):871–899, 2015.

Stamatina Giannouli, Athanasios Kyritsis, Nikolaos Malissovas, Hubert Dominique Becker, and Constantinos Stathopoulos. On the role of an unusual tRNAGly isoacceptor in Staphylococcus aureus. *Biochimie*, 91(3):344–351, 2009.

Carmela M Gibson, T Conn Mallett, Al Claiborne, and Michael G Caparon.

Contribution of NADH oxidase to aerobic metabolism of Streptococcus pyogenes. *Journal of Bacteriology*, 182(2):448–455, 2000.

Richard Giegé, Marie Sissler, and Catherine Florentz. Universal rules and idiosyncratic features in tRNA identity. *Nucleic acids research*, 26(22):5017–5035, 1998.

Nicolas Gisch, Katharina Peters, Ulrich Zhringer, and Waldemar Vollmer. Chapter 8 - the pneumococcal cell wall. In Jeremy Brown, Sven Hammerschmidt, and Carlos Orihuela, editors, *Streptococcus Pneumoniae*, pages 145 – 167. Academic Press, Amsterdam, 2015.

Marjolein Glas, Stephen H McLaughlin, Winfried Roseboom, Fan Liu, Gregory M Koningstein, Alexander Fish, Tanneke den Blaauwen, Albert JR Heck, Luitzen de Jong, Wilbert Bitter, *et al.* The soluble periplasmic domains of Escherichia coli cell division proteins FtsQ/FtsB/FtsL form a trimeric complex with submicromolar affinity. *Journal of Biological Chemistry*, 290(35):21498–21509, 2015.

David S Goodsell and Arthur J Olson. Structural symmetry and protein function. *Annual review of biophysics and biomolecular structure*, 29(1):105–153, 2000.

Kate Gould. Antibiotics: from prehistory to the present day. *Journal of Antimicrobial Chemotherapy*, 71(3):572–575, 02 2016.

Richard J Gowers, Max Linke, Jonathan Barnoud, Tyler JE Reddy, Manuel N Melo, Sean L Seyler, Jan Domański, David L Dotson, Sébastien Buchoux, Ian M Kenney, *et al.* MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. In *Proceedings of the 15th Python in Science Conference*, volume 98. SciPy Austin, TX, 2016.

James A Graham, Jonathan W Essex, and Syma Khalid. PyCGTOOL: automated generation of coarse-grained molecular dynamics models from atomistic trajectories. *Journal of chemical information and modeling*, 57(4):650–656, 2017.

The International Pneumococcal Study Group, S. R. J. Alanee, J. Rello, C. M. Luna,

L. McGee, K. P. Klugman, D. Jackson, L. M. Baddour, V. L. Yu, C. C. Chiou, M. Ip, C. Feldman, A. Ortqvist, and A. J. Morris. Association of Serotypes of Streptococcus pneumoniae with Disease Severity and Outcome in Adults: An International Study. *Clinical Infectious Diseases*, 45(1):46–51, 07 2007.

Alexey Gurevich, Vladislav Saveliev, Nikolay Vyahhi, and Glenn Tesler. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8):1072–1075, 2013.

Regine Hakenbeck, Nadège Balmelle, Beate Weber, Christophe Gardès, Wolfgang Keck, and Antoine de Saizieu. Mosaic Genes and Mosaic Chromosomes: Intra- and Interspecies Genomic Variation of Streptococcus pneumoniae. *Infection and Immunity*, 69(4):2477–2486, 2001.

Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek, and Geoffrey R Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics*, 4(1):17, 2012.

Tsunemi Hasegawa, Masaaki Miyano, Hyouta Himeno, Yoichi Sano, Koichi Kimura, and Mikio Shimizu. Identity determinants of E. coli threonine tRNA. *Biochemical and biophysical research communications*, 184(1):478–484, 1992.

Subray S Hegde and John S Blanchard. Kinetic and mechanistic characterization of recombinant Lactobacillus viridescens FemX (UDP-N-acetylmuramoyl pentapeptide-lysine N6-alanyltransferase). *Journal of Biological Chemistry*, 278 (25):22861–22867, 2003.

Jorgen Henrichsen. Six newly recognized types of Streptococcus pneumoniae. *Journal of clinical microbiology*, 33(10):2759, 1995.

Robert O Heuckeroth, Luis Glaser, and Jeffrey I Gordon. Heteroatom-substituted fatty acid analogs as substrates for N-myristoyltransferase: an approach for

studying both the enzymology and function of protein acylation. *Proceedings of the National Academy of Sciences*, 85(23):8795–8799, 1988.

Adam Hospital, Josep Ramon Goñi, Modesto Orozco, and Josep L Gelpí. Molecular dynamics simulations: advances and applications. *Advances and applications in bioinformatics and chemistry: AABC*, 8:37, 2015.

Ya-Ming Hou and Paul Schimmel. A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature*, 333(6169):140, 1988.

Jing Huang, Sarah Rauscher, Grzegorz Nawrocki, Ting Ran, Michael Feig, Bert L de Groot, Helmut Grubmüller, and Alexander D MacKerell Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods*, 14(1):71, 2017.

Teng-Yi Huang, Jian Liu, and Scott A McLuckey. Top-down tandem mass spectrometry of tRNA via ion trap collision-induced dissociation. *Journal of the American Society for Mass Spectrometry*, 21(6):890–898, 2010.

CC Hyde, SA Ahmed, EA Padlan, E Wilson Miles, and DR Davies. Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium. *Journal of Biological Chemistry*, 263(33):17857–17871, 1988.

Michael Ibba and Dieter Söll. Aminoacyl-tRNA synthesis. *Annual review of biochemistry*, 69(1):617–650, 2000.

Michael Ibba, Kwang-Won Hong, Joyce M Sherman, Sanja Sever, and Dieter Söll. Interactions between tRNA identity nucleotides and their recognition sites in glutaminyl-tRNA synthetase determine the cognate amino acid affinity of the enzyme. *Proceedings of the National Academy of Sciences*, 93(14):6953–6958, 1996.

James A Imlay. Pathways of Oxidative Damage. *Annual Review of Microbiology*, 57 (1):395–418, 2003.

Matthias Imöhl, Ralf René Reinert, Christina Ocklenburg, and Mark van der Linden.

Association of Serotypes of Streptococcus pneumoniae with Age in Invasive Pneumococcal Disease. *Journal of Clinical Microbiology*, 48(4):1291–1296, 2010.

Yuzuru Itoh, Shun-ichi Sekine, Shiro Suetsugu, and Shigeyuki Yokoyama. Tertiary structure of bacterial selenocysteine tRNA. *Nucleic Acids Research*, 41(13): 6729–6738, 2013.

Michal Jamroz, Andrzej Kolinski, and Sebastian Kmiecik. CABS-flex: server for fast simulation of protein structure fluctuations. *Nucleic Acids Research*, 41(W1): W427–W431, 2013.

Jaeho Jeong, Yongsik Jung, Seungjin Na, Jihye Jeong, Eunsun Lee, Mi-Sun Kim, Sun Choi, Dong-Hae Shin, Eunok Paek, Hee-Yoon Lee, *et al.* Novel oxidative modifications in redox-active cysteine residues. *Molecular & cellular proteomics*, 10 (3):M110–000513, 2011.

JE Jimenez-Roldan, RB Freedman, RA Römer, and Stephen A Wells. Rapid simulation of protein motion: merging flexibility, rigidity and normal mode analyses. *Physical biology*, 9(1):016008, 2012.

Sunhwan Jo, Xi Cheng, Jumin Lee, Seonghoon Kim, Sang-Jun Park, Dhilon S Patel, Andrew H Beaven, Kyu Il Lee, Huan Rui, Soohyung Park, *et al.* CHARMM-GUI 10 years for biomolecular modeling and simulation. *Journal of computational chemistry*, 38(15):1114–1124, 2017.

Joanne E Johnson and Rosemary B Cornell. Amphitropic proteins: regulation by reversible membrane interactions. *Molecular membrane biology*, 16(3):217–235, 1999.

John Josse. Constitutive inorganic pyrophosphatase of Escherichia coli I. Purification and catalytic properties. *Journal of Biological Chemistry*, 241(9):1938–1947, 1966.

Garima Kapoor, Saurabh Saigal, and Ashok Elongavan. Action and resistance mechanisms of antibiotics: A guide for clinicians. *Journal of anaesthesiology, clinical pharmacology*, 33(3):300, 2017.

Assaf Katz, Sara Elgamal, Andrei Rajkovic, and Michael Ibba. Non-canonical roles

of tRNAs and tRNA mimics in bacterial cell biology. *Molecular microbiology*, 101 (4):545–558, 2016.

Keith P. Klugman, Kimberly A. Buie, Anne von Gottberg, Shabir A. Madhi, Heather H. Crewe-Brown, Olga Perovic, Alan Karstaedt, and Charles Feldman. Gender as a Risk Factor for Both Antibiotic Resistance and Infection with Pediatric Serogroups/Serotypes, in HIV-Infected and -Uninfected Adults with Pneumococcal Bacteremia. *The Journal of Infectious Diseases*, 189(11):1996–2000, 06 2004.

Sebastian Kmiecik, Dominik Gront, Michal Kolinski, Lukasz Wieteska, Aleksandra Elzbieta Dawid, and Andrzej Kolinski. Coarse-grained protein models and their applications. *Chemical reviews*, 116(14):7898–7936, 2016.

Emily Kramer, Haritha Vallabhaneni, Lauren M Mayer, , and Philip J Farabaugh. A comprehensive analysis of translational missense errors in the yeast Saccharomyces cerevisiae. *RNA*, 16(9):1797–1808, 2010.

Elmar Krieger, Sander B Nabuurs, and Gert Vriend. Homology modeling. *Methods of biochemical analysis*, 44:509–524, 2003.

Alvin CY Kuk, Ellene H Mashalidis, and Seok-Yong Lee. Crystal structure of the MOP flippase MurJ in an inward-facing conformation. *Nature structural & molecular biology*, 24(2):171, 2017.

Sujeet Kumar, Frederick A Rubino, Alicia G Mendoza, and Natividad Ruiz. The bacterial lipid II flippase MurJ functions by an alternating-access mechanism. *Journal of Biological Chemistry*, 294(3):981–990, 2019.

C. G. Kurland. Translational Accuracy and the Fitness of Bacteria. *Annual Review of Genetics*, 26(1):29–50, 1992. doi: 10.1146/annurev.ge.26.120192.000333. URL `https://doi.org/10.1146/annurev.ge.26.120192.000333`. PMID: 1482115.

Ulrich K Laemmli. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *nature*, 227(5259):680, 1970.

Thomas R. Larson and Janet Yother. Streptococcus pneumoniae capsular polysaccharide is linked to peptidoglycan via a direct glycosidic bond to $\beta$-D-N-acetylglucosamine. *Proceedings of the National Academy of Sciences*, 114 (22):5695–5700, 2017. doi: 10.1073/pnas.1620431114.

Gloria Levicán, Assaf Katz, Patricio Valenzuela, Dieter Söll, and Omar Orellana. A tRNAGlu that uncouples protein and tetrapyrrole biosynthesis. *FEBS letters*, 579 (28):6383–6387, 2005.

Michael Levitt. Accurate modeling of protein conformation by automatic segment matching. *Journal of molecular biology*, 226(2):507–533, 1992.

Catherine A Lexau, Ruth Lynfield, Richard Danila, Tamara Pilishvili, Richard Facklam, Monica M Farley, Lee H Harrison, William Schaffner, Arthur Reingold, Nancy M Bennett, *et al.* Changing epidemiology of invasive pneumococcal disease among older adults in the era of pediatric pneumococcal conjugate vaccine. *Jama*, 294(16):2043–2051, 2005.

Rongzhong Li, Lindsay Macnamara, Jessica Leuchter, Rebecca Alexander, and Samuel Cho. MD simulations of tRNA and aminoacyl-tRNA synthetases: dynamics, folding, binding, and allostery. *International journal of molecular sciences*, 16(7):15872–15902, 2015.

Ti-Yu Lin, William S Gross, George K Auer, and Douglas B Weibel. Cardiolipin Alters Rhodobacter sphaeroides Cell Shape by Affecting Peptidoglycan Precursor Biosynthesis. *mBio*, 10(1):e02401–18, 2019.

J. F. Linares, I. Gustafsson, F. Baquero, and J. L. Martinez. Antibiotics as intermicrobial signaling agents instead of weapons. *Proceedings of the National Academy of Sciences*, 103(51):19484–19489, 2006.

Jiqiang Ling and Dieter Söll. Severe oxidative stress induces protein mistranslation through impairment of an aminoacyl-tRNA synthetase editing site. *Proceedings of the National Academy of Sciences*, 107(9):4028–4033, 2010.

Jiqiang Ling, Kaitlyn M Peterson, Ivana Simonović, Dieter Söll, and Miljan Simonović. The mechanism of pre-transfer editing in yeast mitochondrial threonyl-tRNA synthetase. *Journal of Biological Chemistry*, 287(34):28518–28525, 2012.

Richard SA Lipman, Penny J Beuning, Karin Musier-Forsyth, and Ya-Ming Hou. Amino acid activation of a dual-specificity tRNA synthetase is independent of tRNA. *Journal of molecular biology*, 316(3):421–427, 2002.

Xue Liu, Clement Gallay, Morten Kjos, Arnau Domenech, Jelle Slager, Sebastiaan P van Kessel, Kèvin Knoops, Robin A Sorg, Jing-Ren Zhang, and Jan-Willem Veening. High-throughput CRISPRi phenotyping identifies new essential genes in Streptococcus pneumoniae. *Molecular systems biology*, 13(5):931, 2017.

Adrian J Lloyd, Andrea M Gilbey, Anne M Blewett, Gianfranco De Pascale, Ahmed El Zoeiby, Roger C Levesque, Anita C Catherwood, Alexander Tomasz, Timothy DH Bugg, David I Roper, *et al.* Characterization of tRNA-dependent peptide bond formation by MurM in the synthesis of Streptococcus pneumoniae peptidoglycan. *Journal of Biological Chemistry*, 283(10):6402–6417, 2008.

Harvey Lodish, James E Darnell, Arnold Berk, Chris A Kaiser, Monty Krieger, Matthew P Scott, Anthony Bretscher, Hidde Ploegh, Paul Matsudaira, *et al. Molecular cell biology*. Macmillan, 2008.

DT Logan, MH Mazauric, D Kern, and D Moras. Crystal structure of glycyl-tRNA synthetase from Thermus thermophilus. *The EMBO Journal*, 14(17):4156–4167, 1995.

Simon C Lovell, Ian W Davis, W Bryan Arendall III, Paul IW De Bakker, J Michael Word, Michael G Prisant, Jane S Richardson, and David C Richardson. Structure validation by C$\alpha$ geometry: $\phi$, $\psi$ and C$\beta$ deviation. *Proteins: Structure, Function, and Bioinformatics*, 50(3):437–450, 2003.

Nicholas M Luscombe, Susan E Austin, Helen M Berman, and Janet M Thornton.

An overview of the structures of protein-DNA complexes. *Genome biology*, 1(1): reviews001–1, 2000.

Fábio Madeira, Joon Lee, Nicola Buso, Tamer Gur, Nandana Madhusoodanan, Prasad Basutkar, Adrian Tivey, Simon C Potter, Robert D Finn, Rodrigo Lopez, *et al.* The embl-ebi search and sequence analysis tools apis in 2019. *Nucleic acids research*, 2019.

Antoine P Maillard, Sabrina Biarrotte-Sorin, Rgis Villet, Stphane Mesnage, Ahmed Bouhss, Wladimir Sougakoff, Claudine Mayer, and Michel Arthur. Structure-Based Site-Directed Mutagenesis of the UDP-MurNAc-Pentapeptide-Binding Cavity of the FemX Alanyl Transferase from Weissella viridescens. *Journal of Bacteriology*, 187(11):3833–3838, 2005. doi: 10.1128/JB.187.11.3833-3838.2005.

Neelan J Marianayagam, Margaret Sunde, and Jacqueline M Matthews. The power of two: protein dimerization in biology. *Trends in biochemical sciences*, 29(11): 618–625, 2004.

Marc A Martí-Renom, Ashley C Stuart, András Fiser, Roberto Sánchez, Francisco Melo, and Andrej Šali. Comparative protein structure modeling of genes and genomes. *Annual review of biophysics and biomolecular structure*, 29(1):291–325, 2000.

Alexander J Meeske, Eammon P Riley, William P Robins, Tsuyoshi Uehara, John J Mekalanos, Daniel Kahne, Suzanne Walker, Andrew C Kruse, Thomas G Bernhardt, and David Z Rudner. SEDS proteins are a widespread family of bacterial cell wall polymerases. *Nature*, 537(7622):634, 2016.

Susana Mendez-Lage, Isabel Losada-Castillo, and Andres Agulla-Budino. Streptococcus pneumoniae: serotype distribution, antimicrobial susceptibility, risk factors and mortality in Galicia over a two year-period. *Enfermedades infecciosas y microbiologia clinica*, 33(9):579–584, 2015.

Surabhi Mishra and James Imlay. Why do bacteria use so many enzymes to scavenge

hydrogen peroxide? *Archives of Biochemistry and Biophysics*, 525(2):145 – 160, 2012.

Kyle Mohler and Michael Ibba. Translational fidelity and mistranslation in the cellular response to stress. *Nature microbiology*, 2(9):17117, 2017.

Cécile Morlot, Daniel Straume, Katharina Peters, Olav A Hegnar, Nolwenn Simon, Anne-Marie Villard, Carlos Contreras-Martel, Francisco Leisico, Eefjan Breukink, Christine Gravier-Pelletier, *et al.* Structure of the essential peptidoglycan amidotransferase MurT/GatD complex from Streptococcus pneumoniae. *Nature communications*, 9(1):3180, 2018.

Garrett M Morris, Ruth Huey, William Lindstrom, Michel F Sanner, Richard K Belew, David S Goodsell, and Arthur J Olson. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16):2785–2791, 2009.

Lidia Mosyak, Ludmila Reshetnikova, Yehuda Goldgur, Marc Delarue, and Mark G Safro. Structure of phenylalanyl-tRNA synthetase from Thermus thermophilus. *Nature Structural and Molecular Biology*, 2(7):537, 1995.

Daniela Munch, Terry Roemer, Sang Ho Lee, Marianne Engeser, Hans Georg Sahl, and Tanja. Schneider. Identification and in vitro Analysis of the GatD/MurT Enzyme-Complex Catalyzing Lipid II Amidation in Staphylococcus aureus. *PLOS Pathogens*, 8(1):1–11, 01 2012.

Daniel M. Musher, Mark E. Dowell, Virginia D. Shortridge, Robert K. Flamm, James H. Jorgensen, Pierre Le Magueres, and Kurt L. Krause. Emergence of Macrolide Resistance during Treatment of Pneumococcal Pneumonia. *New England Journal of Medicine*, 346(8):630–631, 2002.

Masahiro Naganuma, Shun-ichi Sekine, Ryuya Fukunaga, and Shigeyuki Yokoyama. Unique protein architecture of alanyl-tRNA synthetase for aminoacylation, editing, and dimerization. *Proceedings of the National Academy of Sciences*, 106 (21):8489–8494, 2009.

Tyzoon K Nomanbhoy and Paul R Schimmel. Misactivated amino acids translocate at similar rates across surface of a tRNA synthetase. *Proceedings of the National Academy of Sciences*, 97(10):5119–5122, 2000.

Eva Maria Novoa and Lluís Ribas de Pouplana. Speeding with control: codon usage, tRNAs, and ribosomes. *Trends in Genetics*, 28(11):574–581, 2012.

O Nureki, T Kohno, K Sakamoto, T Miyazawa, and S Yokoyama. Chemical modification and mutagenesis studies on zinc binding of aminoacyl-tRNA synthetases. *Journal of Biological Chemistry*, 268(21):15368–15373, 1993.

Katherine L O'Brien, Lara J Wolfson, James P Watt, Emily Henkle, Maria Deloria-Knoll, Natalie McCall, Ellen Lee, Kim Mulholland, Orin S Levine, Thomas Cherian, *et al.* Burden of disease caused by streptococcus pneumoniae in children younger than 5 years: global estimates. *The Lancet*, 374(9693):893–902, 2009.

Johnsborg Ola, Vegard Eldholm, and Leiv Sigve Hvarstein. Natural genetic transformation: prevalence, mechanisms and function. *Research in Microbiology*, 158(10):767 – 778, 2007.

Jim O'Neill. Review on antimicrobial resistance. *Pharmacy and therapeutics*, 2016.

L Pallanck, Shihong Li, and LH Schulman. The anticodon and discriminator base are major determinants of cysteine tRNA identity in vivo. *Journal of Biological Chemistry*, 267(11):7221–7223, 1992.

Yan Ling Joy Pang, Kiranmai Poruri, and Susan A Martinis. tRNA synthetase: tRNA aminoacylation and beyond. *Wiley Interdisciplinary Reviews: RNA*, 5(4):461–480, 2014.

In Ho Park, David G Pritchard, Rob Cartee, Angela Brandao, Maria Cristina C Brandileone, and Moon H Nahm. Discovery of a new capsular serotype (6C) within serogroup 6 of Streptococcus pneumoniae. *Journal of clinical microbiology*, 45(4):1225–1233, 2007.

Carol E Parker, Viorel Mocanu, Mihaela Mocanu, Nedyalka Dicheva, and Maria R

Warren. Mass spectrometry for post-translational modifications. *Neuroproteomics*, 2010:PMID–21882444, 2010.

Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

Christopher D Pericone, Karin Overweg, Peter W.M Hermans, and Jeffrey N. Weiser. Inhibitory and Bactericidal Effects of Hydrogen Peroxide Production by Streptococcus pneumoniae on Other Inhabitants of the Upper Respiratory Tract. *Infection and Immunity*, 68(7):3990–3997, 2000.

Christopher D Pericone, Sunny Park, James A Imlay, and Jeffrey N. Weiser. Factors Contributing to Hydrogen Peroxide Resistance in Streptococcus pneumoniae Include Pyruvate Oxidase (SpxB) and Avoidance of the Toxic Effects of the Fenton Reaction. *Journal of Bacteriology*, 185(23):68156825, 2003.

John J Perona and Ita Gruic-Sovulj. Synthetic and editing mechanisms of aminoacyl-tRNA synthetases. In *Aminoacyl-tRNA Synthetases in Biology and Medicine*, pages 1–41. Springer, 2013.

Brian G. Pierce, Kevin Wiehe, Howook Hwang, Bong-Hyun Kim, Thom Vreven, and Zhiping Weng. ZDOCK server: interactive docking prediction of proteinprotein complexes and symmetric multimers. *Bioinformatics*, 30(12):1771–1773, 2014.

Roswell Quinn. Rethinking antibiotic research and development: World War II and the penicillin collaborative. *American journal of public health*, 103(3):426–434, 2013.

Venki Ramakrishnan. Ribosome structure and the mechanism of translation. *Cell*, 108(4):557–572, 2002.

Allan Rauch, Angela B. Brueggemann, Gary V. Doern, Jr. Holley, H. Preston, Jen Duval, Karen Sue Kehl, Kari Kugler, Michael Blocker, Michael Dunne, Michael A. Pfaller, and Shannon Putnam. Clonal Relationships Among High-Level Penicillin-Resistant Streptococcus pneumoniae in the United States. *Clinical*

*Infectious Diseases*, 27(4):757–761, 10 1998. ISSN 1058-4838. doi: 10.1086/514937. URL https://doi.org/10.1086/514937.

Gili Regev-Yochay, Krzysztof Trzcinski, Claudette M Thompson, Marc Lipsitch, and Richard Malley. SpxB is a suicide gene of Streptococcus pneumoniae and confers a selective advantage in an in vivo competitive colonization model. *Journal of bacteriology*, 189(18):6532–6539, 2007.

R.R. Reinert. The antimicrobial resistance profile of Streptococcus pneumoniae. *Clinical Microbiology and Infection*, 15:7 – 11, 2009.

Susanne Rohrer and Brigitte Berger-Bchi. Application of a bacterial two-hybrid system for the analysis of proteinprotein interactions between FemABX family proteins. *Microbiology*, 149(10):2733–2738, 2003.

Rudolf A Römer, Stephen A Wells, J Emilio Jimenez-Roldan, Moitrayee Bhattacharyya, Saraswathi Vishweshwara, and Robert B Freedman. The flexibility and dynamics of protein disulfide isomerase. *Proteins: Structure, Function, and Bioinformatics*, 84(12):1776–1785, 2016.

Ambrish Roy, Alper Kucukural, and Yang Zhang. I-tasser: a unified platform for automated protein structure and function prediction. *Nature protocols*, 5(4):725, 2010.

Natividad. Ruiz. Lipid flippases for bacterial peptidoglycan biosynthesis. *Lipid Insights*, 8s1:LPI.S31783, 2015. doi: 10.4137/LPI.S31783. URL https://doi.org/10.4137/LPI.S31783.

Andrej Šali and Tom L Blundell. Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology*, 234(3):779–815, 1993.

Roberto Sanchez and Andrej Šali. Advances in comparative protein-structure modelling. *Current opinion in structural biology*, 7(2):206–214, 1997.

Lee E Sanderson and Olke C. Uhlenbeck. The 51-63 base pair of tRNA confers specificity for binding by EF-Tu. 13:835–40, 07 2007.

Eric Sauvage, Frdric Kerff, Mohammed Terrak, Juan A. Ayala, and Paulette Charlier. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiology Reviews*, 32(2):234–258, 02 2008.

Albert Schatz, Elizabeth Bugle, and Selman A Waksman. Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria. *Proceedings of the Society for Experimental Biology and Medicine*, 55(1): 66–69, 1944.

Dirk-Jan Scheffers and Mariana G Pinho. Bacterial cell wall synthesis: new insights from localization studies. *Microbiology and Molecular Biology Reviews*, 69(4):585–607, 2005.

Paul Schimmel. Classes of aminoacyl-tRNA synthetases and the establishment of the genetic code. *Trends in biochemical sciences*, 16:1–3, 1991.

Paul R. Schimmel and Dieter Söll. Aminoacyl-tRNA Synthetases: General features and recognition of transfer RNAs. *Annual Review of Biochemistry*, 48(1):601–648, 1979.

T Martin Schmeing, Rebecca M Voorhees, Ann C Kelley, Yong-Gui Gao, Frank V Murphy, John R Weir, and V Ramakrishnan. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science*, 326(5953):688–694, 2009.

Tanja Schneider and Hans-Georg Sahl. An oldie but a goodie cell wall biosynthesis as antibiotic target pathway. *International Journal of Medical Microbiology*, 300(2): 161 – 169, 2010.

Tanja Schneider, Maria M. Senn, Brigitte BergerBchi, Alessandro Tossi, HansGeorg Sahl, and Imke Wiedemann. In vitro assembly of a complete, pentaglycine interpeptide bridge containing cell wall precursor (lipid IIGly5) of Staphylococcus aureus.

Torsten Schwede. Protein modeling: what happened to the protein structure gap? *Structure*, 21(9):1531–1540, 2013.

Torsten Schwede, Andrej Sali, Barry Honig, Michael Levitt, Helen M Berman, David Jones, Steven E Brenner, Stephen K Burley, Rhiju Das, Nikolay V Dokholyan, *et al.* Outcome of a workshop on applications of protein models in biomedical research. *Structure*, 17(2):151–159, 2009.

Torsten Seemann. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30 (14):2068–2069, 2014.

Masanori Seki, Ken-ichiro Iida, Mitsumasa Saito, Hiroaki Nakayama, and Shin-ichi Yoshida. Hydrogen Peroxide Production in Streptococcus pyogenes: Involvement of Lactate Oxidase and Coupling with Aerobic Utilization of Lactate. *Journal of Bacteriology*, 186(7):2046–2051, 2004. doi: 10.1128/JB.186.7.2046-2051.2004.

A Severin and A Tomasz. Naturally occurring peptidoglycan variants of Streptococcus pneumoniae. *Journal of bacteriology*, 178(1):168–174, 1996.

Samir S Shah and Adam J Ratner. Trends in invasive pneumococcal diseaseassociated hospitalizations. *Clinical Infectious Diseases*, 42(1):e1–5, 2006.

Md Munan Shaik, Amandine Maccagni, Guillaume Tourcier, Anne Marie Di Guilmi, and Andréa Dessen. Structural basis of pilus anchoring by the ancillary pilin RrgC of Streptococcus pneumoniae. *Journal of Biological Chemistry*, 289(24):16988–16997, 2014.

Min-yi Shen and Andrej Sali. Statistical potential for assessment and prediction of protein structures. *Protein science*, 15(11):2507–2524, 2006.

Jennifer Shepherd. Characterisation of Pneumococcal Peptidoglycan Cross-linking Enzymology. 2011.

Jennifer Shepherd and Michael Ibba. Lipid II-Independent trans Editing of Mischarged tRNAs by the Penicillin Resistance Factor MurM. *The Journal of Biological Chemistry*, 288(36):25915–25923, 2013a.

Jennifer Shepherd and Michael Ibba. Direction of aminoacylated transfer RNAs

into antibiotic synthesis and peptidoglycan-mediated antibiotic resistance. *FEBS letters*, 587(18):2895–904, sep 2013b.

Jennifer Shepherd and Michael Ibba. Relaxed substrate specificity leads to extensive tRNA mischarging by Streptococcus pneumoniae class I and class II aminoacyl-tRNA synthetases. *MBio*, 5(5):e01656–14, 2014.

Huijing Shi and Peter B Moore. The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *Rna*, 6(8):1091–1105, 2000.

Fabian Sievers, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, Johannes Söding, *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, 7(1):539, 2011.

Lynn L Silver. Challenges of antibacterial discovery. *Clinical microbiology reviews*, 24 (1):71–109, 2011.

Anthony M Smith and Keith P Klugman. Non-penicillin-binding protein mediated high-level penicillin and cephalosporin resistance in a Hungarian clone of Streptococcus pneumoniae. *Microbial Drug Resistance*, 6(2):105–110, 2000.

Anthony M. Smith and Keith P. Klugman. Alterations in MurM, a Cell Wall Muropeptide Branching Enzyme, Increase High-Level Penicillin and Cephalosporin Resistance in Streptococcus Pneumoniae. *Antimicrobial Agents and Chemotherapy*, 45(8):23932396, 2001.

MS Smyth and JHJ Martin. x ray crystallography. *Molecular Pathology*, 53(1):8, 2000.

Barbara Spellerberg, Diana R Cundell, Jens Sandros, Barbara J Pearce, Ilona Idänpään-Heikkilä, Carsten Rosenow, and H Robert Masure. Pyruvate oxidase, as a determinant of virulence in Streptococcus pneumoniae. *Molecular microbiology*, 19(4):803–813, 1996.

Mathias Sprinzl, Thomas Hartmann, Joachim Weber, Jutta Blank, and Robert Zeidler.

Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic acids research*, 17(Suppl):r1, 1989.

F William Studier and Barbara A Moffatt. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology*, 189(1):113–130, 1986.

Atsushi Taguchi, Michael A Welsh, Lindsey S Marmont, Wonsik Lee, Megan Sjodt, Andrew C Kruse, Daniel Kahne, Thomas G Bernhardt, and Suzanne Walker. FtsW is a peptidoglycan polymerase that is functional only in complex with its cognate penicillin-binding protein. *Nature microbiology*, page 1, 2019.

Iida Ken-ichiro. Seki Masanori. Saito Mitsumasa. Shiota Susumu. Nakayama Hiroaki Taniai, Hiroaki. and Shin-ichi. Yoshida. Concerted Action of Lactate Oxidase and Pyruvate Oxidase in Aerobic Growth of Streptococcus Pneumoniae: Role of Lactate as an Energy Source. *Journal of Bacteriology*, 190(10):35723579, 2008.

Maret du Toit, Melanie Huch, Gyu-Sung Cho, and Charles MAP Franz. The genus Streptococcus. *Lactic Acid Bacteria: Biodiversity and Taxonomy*, pages 457–505, 2014.

Huichun Tong, Xinhui Wang, Yuzhu Dong, Qingqing Hu, Ziyi Zhao, Yun Zhu, Linxuan Dong, Fan Bai, and Xiuzhu Dong. A Streptococcus aquaporin acts as peroxiporin for efflux of cellular hydrogen peroxide and alleviation of oxidative stress. *Journal of Biological Chemistry*, pages jbc–RA118, 2019.

Henrique Tono and Arthur Kornberg. Biochemical studies of bacterial sporulation III. Inorganic pyrophosphatase of vegetative cells and spores of Bacillus subtilis. *Journal of Biological Chemistry*, 242(10):2375–2382, 1967.

Oleg Trott and Arthur J Olson. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2):455–461, 2010.

Athanasios Typas, Manuel Banzhaf, Carol A Gross, and Waldemar Vollmer. From

the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature reviews Microbiology*, 10(2):123136, 2011.

Ilya A Vakser. Protein-protein docking: From interaction to interactome. *Biophysical journal*, 107(8):1785–1793, 2014.

C Lee Ventola. The antibiotic resistance crisis: part 1: causes and threats. *Pharmacy and therapeutics*, 40(4):277, 2015.

Waldemar Vollmer, Didier Blanot, and Miguel A. De Pedro. Peptidoglycan structure and architecture. *FEMS Microbiology Reviews*, 32(2):149–167, 02 2008.

Gunter von Ehrenstein. Isolation of sRNA from intact Escherichia coli cells. *Methods in Enzymology*, 12(A):588–596, 1967.

Selman A Waksman and Hugh B Woodruff. Actinomyces antibioticus, a New Soil Organism Antagonistic to Pathogenic and Non-pathogenic Bacteria. *Journal of bacteriology*, 42 2:231–49, 1941.

Christopher Walsh. Molecular mechanisms that confer antibacterial drug resistance. *Nature*, 406(6797):775, 2000.

Caiyan Wang, Yu Guo, Qingnan Tian, Qian Jia, Yuanzhu Gao, Qinfen Zhang, Chun Zhou, and Wei Xie. SerRS-tRNASec complex structures reveal mechanism of the first step in selenocysteine biosynthesis. *Nucleic acids research*, 43(21):10534–10545, 2015.

Junmei Wang, Romain M Wolf, James W Caldwell, Peter A Kollman, and David A Case. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9):1157–1174, 2004.

Andrew M. Waterhouse, David M. A. Martin, Geoffrey J. Barton, James B. Procter, and Michle Clamp. Jalview Version 2a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9):1189–1191, 01 2009. ISSN 1367-4803. doi: 10.1093/bioinformatics/btp033. URL `https://doi.org/10.1093/bioinformatics/btp033`.

M R. Webb. A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, 89(11):48844887, 1992.

Ian Wheeldon, Shelley D Minteer, Scott Banta, Scott Calabrese Barton, Plamen Atanassov, and Matthew Sigman. Substrate channelling as an approach to cascade reactions. *Nature chemistry*, 8(4):299, 2016.

Cynthia G Whitney, Monica M Farley, James Hadler, Lee H Harrison, Nancy M Bennett, Ruth Lynfield, Arthur Reingold, Paul R Cieslak, Tamara Pilishvili, Delois Jackson, *et al.* Decline in invasive pneumococcal disease after the introduction of protein–polysaccharide conjugate vaccine. *New England Journal of Medicine*, 348 (18):1737–1746, 2003.

Donald E Williams and Ralph A Reisfeld. Disc electrophoresis in polyacrylamide gels: extension to new conditions of pH and buffer. *Annals of the New York Academy of Sciences*, 121(2):373–381, 1964.

Sarah Witzke, Michael Petersen, Timothy S Carpenter, and Syma Khalid. Molecular Dynamics Simulations Reveal the Conformational Flexibility of Lipid II and Its Loose Association with the Defensin Plectasin in the Staphylococcus aureus Membrane. *Biochemistry*, 55(23):3303–3314, 2016.

Megan H Wright, William P Heal, David J Mann, and Edward W Tate. Protein myristoylation in health and disease. *Journal of chemical biology*, 3(1):19–35, 2010.

Jiang Wu, Yongqiang Fan, and Jiqiang Ling. Mechanism of oxidant-induced mistranslation by threonyl-tRNA synthetase. *Nucleic acids research*, 42(10): 6523–6531, 2014.

Ioannis Xenarios, Lukasz Salwinski, Xiaoqun Joyce Duan, Patrick Higney, Sul-Min Kim, and David Eisenberg. DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic acids research*, 30(1):303–305, 2002.

Srujana S Yadavalli and Michael Ibba. Quality control in aminoacyl-tRNA synthesis: its role in translational fidelity. In *Advances in protein chemistry and structural biology*, volume 86, pages 1–43. Elsevier, 2012.

Jianyi Yang, Renxiang Yan, Ambrish Roy, Dong Xu, Jonathan Poisson, and Yang Zhang. The I-TASSER Suite: protein structure and function prediction. *Nature methods*, 12(1):7, 2015.

Hani S Zaher and Rachel Green. Fidelity at the molecular level: lessons from protein synthesis. *Cell*, 136(4):746–762, 2009.

Yang Zhang. I-TASSER server for protein 3D structure prediction. *BMC bioinformatics*, 9(1):40, 2008.

Sanduo Zheng, Lok-To Sham, Frederick A Rubino, Kelly P Brock, William P Robins, John J Mekalanos, Debora S Marks, Thomas G Bernhardt, and Andrew C Kruse. Structure and mutagenic analysis of the lipid II flippase MurJ from Escherichia coli. *Proceedings of the National Academy of Sciences*, 115(26):6709–6714, 2018.

G Zubay. The isolation and fractionation of soluble ribonucleic acid. *Journal of Molecular Biology*, 4(5):347–356, 1962.